CF

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
10 July 2003 (10.07.2003)

PCT

(10) International Publication Number
WO 03/055904 A2

(51) International Patent Classification[7]: C07K 14/195

(21) International Application Number: PCT/CA02/01977

(22) International Filing Date:
20 December 2002 (20.12.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/344,112 21 December 2001 (21.12.2001) US
60/370,755 8 April 2002 (08.04.2002) US

(71) Applicant *(for all designated States except US)*: **AFFINIUM PHARMACEUTICALS, INC.** [CA/CA]; 100 University Avenue, 10th Floor, South Tower, Toronto, Ontario M5J 1V6 (CA).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **EDWARDS, Aled** [CA/CA]; 21 Sutherland Drive, Toronto, Ontario M4G 1H1 (CA). **DHARAMSI, Akil** [CA/CA]; 29 Moresby Street, Richmond Hill, Ontario L4B 4K9 (CA). **DOMAGALA, Megan** [CA/CA]; #135-815 Hillcrest Ave., Mississauga, Ontario L5B 4B1 (CA). **PINDER, Benjamin** [CA/CA]; 181 Brunswick Ave., Apt. 1A, Toronto, Ontario M5S 2M4 (CA). **ALAM, Muhammad, Zahoor** [PK/CA]; #11-440 Malaga Road, Oshawa, Ontario L1J 1N8 (CA). **VEDADI, Masoud** [CA/CA]; 178 Northwood Drive, Toronto, Ontario M2M 2K3 (CA). **WREZEL, Olga** [CA/CA]; The Credit Woodlands, #3514, Mississauga, Ontario L5C 2K6 (CA). **HOUSTON, Simon** [GB/CA]; 100 Vaughan Road, Apt. 29, Toronto, Ontario M6C 2M1 (CA). **KIMBER, Matthew** [GB/CA]; 306 Sumach Street, Apt. 7, Toronto, Ontario M5A 3K2 (CA). **VALLEE, Francois** [FR/CA]; 490 Eglinton Avenue East, Apt. 405, Toronto, Ontario M4P 1M4 (CA). **AWREY, Donald** [CA/CA]; 2211 Stir Crescent, Mississauga, Ontario L4Y 3V2 (CA). **BEATTIE, Bryan** [CA/CA]; 2224 Vista Oak Road, Oakville, Ontario L6M 3L7 (CA).

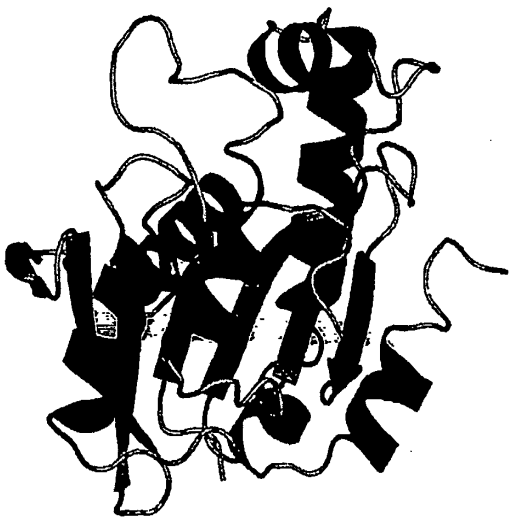
(74) Agent: **BERESKIN & PARR**; 40 King Street West, 40th Floor, Toronto, Ontario M5H 3Y2 (CA).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE,

[Continued on next page]

(54) Title: NOVEL PURIFIED POLYPEPTIDES FROM PSEUDOMONAS AERUGINOSA



(57) Abstract: The present invention relates to novel drug targets for pathogenic bacteria. Accordingly, the invention provides purified protein comprising the amino acid sequence set forth in SEQ ID NO:4. The invention also provides biochemical and biophysical characteristics of the polypeptides of the invention.

WO 03/055904 A2

ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

NOVEL PURIFIED POLYPEPTIDES FROM PSEUDOMONAS AERUGINOSA

RELATED APPLICATION INFORMATION

This application claims the benefit of priority to the following U.S. Provisional Patent Applications, all of which applications are hereby incorporated by reference in their entirety.

Provisional Application Number	*Filing Date*
60/344,112	21-December-2001
60/370,755	08-April-2002

INTRODUCTION

The discovery of novel antimicrobial agents that work by novel mechanisms is a problem researchers in all fields of drug development face today. The increasing prevalence of drug-resistant pathogens (bacteria, fungi, parasites, etc.) has led to significantly higher mortality rates from infectious diseases and currently presents a serious crisis worldwide. Despite the introduction of second and third generation antimicrobial drugs, certain pathogens have developed resistance to all currently available drugs.

One of the problems contributing to the development of multiple drug resistant pathogens is the limited number of protein targets for antimicrobial drugs. Many of the antibiotics currently in use are structurally related or act through common targets or pathways. Accordingly, adaptive mutation of a single gene may render a pathogenic species resistant to multiple classes of antimicrobial drugs. Therefore, the rapid discovery of drug targets is urgently needed in order to combat the constantly evolving threat by such infectious microorganisms.

Recent advances in bacterial and viral genomics research provides an opportunity for rapid progress in the identification of drug targets. The complete genomic sequences for a number of microorganisms are available. However, knowledge of the complete genomic sequence is only the first step in a long process toward discovery of a viable drug target. The genomic sequence must be annotated to identify open reading frames (ORFs), the essentiality of the protein encoded by the ORF must be determined and the mechanism of action of the gene product must be determined in order to develop a targeted approach to drug discovery.

There are a variety of computer programs available to annotate genomic sequences. Genome annotation involves both identification of genes as well assignment of function

thereto based on sequence comparison to homologous proteins with known or predicted functions. However, genome annotation has turned out to be much more of an art than a science. Factors such as splice variants and sequencing errors coupled with the particular algorithms and databases used to annotate the genome can result in significantly different annotations for the same genome. For example, upon reanalysis of the genome of *Mycoplasma pneumoniae* using more rigorous sequence comparisons coupled with molecular biological techniques, such as gel electrophoresis and mass spectrometry, researchers were able to identify several previously unidentified coding sequences, to dismiss a previous identified coding sequence as a likely pseudogene, and to adjust the length of several previously defined ORFs (Dandkar et al. (2000) Nucl. Acids Res. 28(17): 3278-3288). Furthermore, while overall conservation between amino acid sequences generally indicates a conservation of structure and function, specific changes at key residues can lead to significant variation in the biochemical and biophysical properties of a protein. In a comparison of three different functional annotations of the *Mycoplasma genitalium* genome, it was discovered that some genes were assigned three different functions and it was estimated that the overall error rate in the annotations was at least 8% (Brenner (1999) Trends Genet 15(4): 132-3). Accordingly, molecular biological techniques are required to ensure proper genome annotation and identify valid drug targets.

However, confirmation of genome annotation using molecular biological techniques is not an easy proposition due to the unpredictability in expression and purification of polypeptide sequences. Further, in order to carry out structural studies to validate proteins as potential drug targets, it is generally necessary to modify the native proteins in order to facilitate these analyses, e.g., by labeling the protein (e.g., with a heavy atom, isotopic label, polypeptide tag, etc.) or by creating fragments of the polypeptide corresponding to functional domains of a multi-domain protein. Moreover, it is well-known that even small changes in the amino acid sequence of a protein may lead to dramatic affects on protein solubility (Eberstadt et al. (1998) Nature 392: 941-945). Accordingly, genome-wide validation of protein targets will require considerable effort even in light of the sequence of the entire genome of an organism and/or purification conditions for homologs of a particular target.

We have developed reliable, high throughput methods to address some of the shortcomings identified above. In part, using these methods, we have now identified, expressed, and purified a novel antimicrobial target from *Pseudomonas aeruginosa*, or *P.*

aeruginosa. Various biophysical, bioinformatic and biochemical studies have been used to characterize the structure and function of the polypeptides of the invention.

TABLE OF CONTENTS

SUMMARY OF THE INVENTION

As part of an effort at genome-wide structural and functional characterization of microbial targets, the present invention provides polypeptides from *P. aeruginosa*. In various aspects, the invention provides the nucleic acid and amino acid sequences of the polypeptides of the invention. The invention also provides purified, soluble forms of the polypeptides of the invention suitable for structural and functional characterization using a variety of techniques, including, for example, affinity chromatography, mass spectrometry, NMR and x-ray crystallography. The invention further provides modified versions of the polypeptides of the invention to facilitate characterization, including polypeptides labeled with isotopic or heavy atoms and fusion proteins.

A polypeptide of the invention has been crystallized and its structure solved as described in detail below, thereby providing information about the structure of the polypeptide, and druggable regions, domains and the like contained therein, all of which may be used in rational-based drug design efforts.

In general, the biological activity of a polypeptide of the invention is expected to be characterized as having a biochemical activity substantially similar to that of peptidyl-tRNA hydrolase, having the gene designation of *pth*, as described in more detail below.

This assignment has been confirmed by solving the X-ray structure of a polypeptide of the invention.

All of the information learned and described herein about the polypeptides of the invention may be used to design modulators of one or more of their biological activities. In particular, information critical to the design of therapeutic and diagnostic molecules, including, for example, the protein domain, druggable regions, structural information, and the like for the polypeptides of the invention is now available or attainable as a result of the ability to prepare, purify and characterize them, and domains, fragments, variants and derivatives thereof.

In other aspects of the invention, structural and functional information about the polypeptides of the invention has and will be obtained. Such information, for example, may be incorporated into databases containing information on the polypeptides of the invention, as well as other polypeptide targets from other microbial species. Such databases will provide investigators with a powerful tool to analyze the polypeptides of the invention and aid in the rapid discovery and design of therapeutic and diagnostic molecules.

In another aspect, modulators, inhibitors, agonists or antagonists against the polypeptides of the invention, or biological complexes containing them, or orthologues thereto, may be used to treat any disease or other treatable condition of a patient (including humans and animals), and particularly a disease caused by *P. aeruginosa*, such as, for example, one of the following: osteomyelitis, otitis externa, conjunctivitis, keratitis, endophthalmitis, alveolar necrosis, vascular invasion, bacteremia, and burn infection.

The present invention further allows relationships between polypeptides from the same and multiple species to be compared by isolating and studying the various polypeptides of the invention and other proteins. By such comparison studies, which may involve multi-variable analysis as appropriate, it is possible to identify drugs that will affect multiple species or drugs that will affect one or a few species. In such a manner, so-called "wide spectrum" and narrow spectrum" anti-infectives may be identified. Alternatively, drugs that are selective for one or more bacterial or other non-mammalian species, and not for one or more mammalian species (especially human), may be identified (and vice-versa).

In other embodiments, the invention contemplates kits including the subject nucleic acids, polypeptides, crystallized polypeptides, antibodies, and other subject materials, and optionally instructions for their use. Uses for such kits include, for example, diagnostic and therapeutic applications.

The embodiments and practices of the present invention, other embodiments, and their features and characteristics, will be apparent from the description, figures and claims that follow, with all of the claims hereby being incorporated by this reference into this Summary.

5

BRIEF DESCRIPTION OF THE FIGURES

FIGURE 1 shows the nucleic acid coding sequence for an exemplary polypeptide of the invention as predicted from the genomic sequence of *P. aeruginosa* (SEQ ID NO: 1). This predicted nucleic acid coding sequence was cloned and sequenced to produce the 10 polynucleotide sequence shown in FIGURE 2 (SEQ ID NO: 3).

FIGURE 2 shows the amino acid sequence for an exemplary polypeptide of the invention as predicted from the nucleotide sequence shown in FIGURE 1 (SEQ ID NO: 2).

FIGURE 3 shows the experimentally determined nucleic acid coding sequence for an exemplary polypeptide of the invention (SEQ ID NO: 3).

15 FIGURE 4 shows the amino acid sequence for the exemplary polypeptide of the invention as predicted from the nucleotide sequence shown in FIGURE 3 (SEQ ID NO: 4).

FIGURE 5 shows the primer sequences used to amplify the nucleic acid of SEQ ID NO: 3. The primers are SEQ ID NO: 5 and SEQ ID NO: 6.

FIGURE 6 contains Table 1, which provides among other things a variety of data 20 and other information on the polypeptides of the invention.

FIGURE 7 contains Table 2, which provides the results of several bioinformatic analyses relating to SEQ ID NO: 2.

FIGURE 8 contains Table 3, which shows information related to the x-ray structure for a polypeptide of the invention as described more fully in EXAMPLE 16.

25 FIGURE 9 lists the atomic structure coordinates for a polypeptide of the invention derived from x-ray diffraction from a crystal of such polypeptide, as described in more detail in EXAMPLE 16. There are multiple pages to FIGURE 9, labeled 1, 2, 3, etc. The information in such Figure is presented in the following tabular format, with a generic entry provided as an example:

Record Header	No.	Atom Type	Resi -due	Residue Number	X	Y	Z	OCC	B
ATOM 1	1	CB	HIS	1	4.497	15.607	34.172	1	70.54

In the table, "Record Header" describes the row type, such as "ATOM". "No." refers to the row number. The first "Atom Type" column refers to the atom whose coordinates are measured, with the first letter in the column identifying the atom by its elemental symbol and the subsequent letter defining the location of the atom in the amino acid residue or other molecule. "Residue" and "residue number" identifies the residue of the subject polypeptide. "X, Y, Z" crystallographically define the atomic position of the atom measured. "Occ" is an occupancy factor that refers to the fraction of the molecules in which each atom occupies the position specified by the coordinates. A value of "1" indicates that each atom has the same conformation, i.e., the same position, in all molecules of the crystal. "B" is a thermal factor that is related to the root mean square deviation in the position of the atom around the given atomic coordinate.

FIGURE 10 depicts a clustal V-based sequence alignment of the peptidyl-tRNA hydrolase protein sequences from six pathogens. The dark shading indicates conserved amino acids across species, with gray areas less conserved.

FIGURE 11 depicts a ribbon representation of a polypeptide of the invention, peptidyl-tRNA hydrolase from *P. aeruginosa*. The ribbon representation was generated with PYMOL as were all of the following figures (Delano, W.L., The PYMOL Molecular Graphics System (2002) DeLano Scientific, San Carlos, CA, USA). The helices and strands are colored in red and in green, respectively.

FIGURE 12 depicts the superimposition of the *E. coli* (colored in blue) and *P. aeruginosa* (colored in green) crystal structures for peptidyl-tRNA hydrolase.

FIGURE 13 depicts the location of the conserved residues in the three dimensional model of *P. aeruginosa* peptidyl-tRNA hydrolase in the druggable region that is believed to be a peptide-binding active site region, as described in more detail below. The strictly conserved residues among peptidyl-tRNA hydrolase polypeptides from various sources are colored in red. Conservative substitutions are colored in orange and variables residues are colored in yellow. (A) Ribbon diagram topped with side-chains drawn as red, orange and yellow sticks. (B) Molecular surface diagram of a monomer of *P. aeruginosa* peptidyl-tRNA hydrolase shown in roughly the same orientation.

FIGURE 14 depicts close views of what is believed to be the tRNA binding active site channel, as described in detail below. (A) *P. aeruginosa* peptidyl-tRNA hydrolase, with the most conserved residues area colored in red and the less conserved residues area in

blue, based on the data from FIGURE 10. (B) Cross-section of the active site channel. In *E. coli* peptidyl-tRNA hydrolase (on the right), a tryptophan residue partially occludes the mid-section of the channel, as described in greater detail below.

DETAILED DESCRIPTION OF THE INVENTION

1. Definitions

For convenience, certain terms employed in the specification, examples, and appended claims are collected here. Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs.

The articles "a" and "an" are used herein to refer to one or to more than one (i.e., to at least one) of the grammatical object of the article. By way of example, "an element" means one element or more than one element.

The term "amino acid" is intended to embrace all molecules, whether natural or synthetic, which include both an amino functionality and an acid functionality and capable of being included in a polymer of naturally-occurring amino acids. Exemplary amino acids include naturally-occurring amino acids; analogs, derivatives and congeners thereof; amino acid analogs having variant side chains; and all stereoisomers of any of any of the foregoing.

The term "binding" refers to an association, which may be a stable association, between two molecules, e.g., between a polypeptide of the invention and a binding partner, due to, for example, electrostatic, hydrophobic, ionic and/or hydrogen-bond interactions under physiological conditions.

A "comparison window," as used herein, refers to a conceptual segment of at least 20 contiguous amino acid positions wherein a protein sequence may be compared to a reference sequence of at least 20 contiguous amino acids and wherein the portion of the protein sequence in the comparison window may comprise additions or deletions (i.e., gaps) of 20 percent or less as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. Optimal alignment of sequences for aligning a comparison window may be conducted by the local homology algorithm of Smith and Waterman (1981) Adv. Appl. Math. 2: 482, by the homology alignment algorithm of Needleman and Wunsch (1970) J. Mol. Biol. 48: 443, by the search for similarity method of Pearson and Lipman (1988) Proc. Natl. Acad. Sci. (U.S.A.) 85:

2444, by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package Release 7.0, Genetics Computer Group, 575 Science Dr., Madison, WI), or by inspection, and the best alignment (i.e., resulting in the highest percentage of homology over the comparison window) generated by the various methods may be identified.

The term "complex" refers to an association between at least two moieties (e.g. chemical or biochemical) that have an affinity for one another. Examples of complexes include associations between antigen/antibodies, lectin/avidin, target polynucleotide/probe oligonucleotide, antibody/anti-antibody, receptor/ligand, enzyme/ligand, polypeptide/polypeptide, polypeptide/polynucleotide, polypeptide/co-factor, polypeptide/substrate, polypeptide/inhibitor, polypeptide/small molecule, and the like. "Member of a complex" refers to one moiety of the complex, such as an antigen or ligand. "Protein complex" or "polypeptide complex" refers to a complex comprising at least one polypeptide.

The term "conserved residue" refers to an amino acid that is a member of a group of amino acids having certain common properties. The term "conservative amino acid substitution" refers to the substitution (conceptually or otherwise) of an amino acid from one such group with a different amino acid from the same group. A functional way to define common properties between individual amino acids is to analyze the normalized frequencies of amino acid changes between corresponding proteins of homologous organisms (Schulz, G. E. and R. H. Schirmer., Principles of Protein Structure, Springer-Verlag). According to such analyses, groups of amino acids may be defined where amino acids within a group exchange preferentially with each other, and therefore resemble each other most in their impact on the overall protein structure (Schulz, G. E. and R. H. Schirmer, Principles of Protein Structure, Springer-Verlag). One example of a set of amino acid groups defined in this manner include: (i) a charged group, consisting of Glu and Asp, Lys, Arg and His, (ii) a positively-charged group, consisting of Lys, Arg and His, (iii) a negatively-charged group, consisting of Glu and Asp, (iv) an aromatic group, consisting of Phe, Tyr and Trp, (v) a nitrogen ring group, consisting of His and Trp, (vi) a large aliphatic nonpolar group, consisting of Val, Leu and Ile, (vii) a slightly-polar group, consisting of Met and Cys, (viii) a small-residue group, consisting of Ser, Thr, Asp, Asn, Gly, Ala, Glu, Gln and Pro, (ix) an aliphatic group consisting of Val, Leu, Ile, Met and Cys, and (x) a small hydroxyl group consisting of Ser and Thr.

The term "domain", when used in connection with a polypeptide, refers to a specific region within such polypeptide that comprises a particular structure or mediates a particular function. In the typical case, a domain of a polypeptide of the invention is a fragment of the polypeptide. In certain instances, a domain is a structurally stable domain, as evidenced, for example, by mass spectroscopy, or by the fact that a modulator may bind to a druggable region of the domain.

The term "druggable region", when used in reference to a polypeptide, nucleic acid, complex and the like, refers to a region of the molecule which is a target or is a likely target for binding a modulator. For a polypeptide, a druggable region generally refers to a region wherein several amino acids of a polypeptide would be capable of interacting with a modulator or other molecule. For a polypeptide or complex thereof, exemplary druggable regions including binding pockets and sites, enzymatic active sites, interfaces between domains of a polypeptide or complex, surface grooves or contours or surfaces of a polypeptide or complex which are capable of participating in interactions with another molecule. In certain instances, the interacting molecule is another polypeptide, which may be naturally-occurring. In other instances, the druggable region is on the surface of the molecule.

Druggable regions may be described and characterized in a number of ways. For example, a druggable region may be characterized by some or all of the amino acids that make up the region, or the backbone atoms thereof, or the side chain atoms thereof (optionally with or without the Cα atoms). Alternatively, in certain instances, the volume of a druggable region corresponds to that of a carbon based molecule of at least about 200 amu and often up to about 800 amu. In other instances, it will be appreciated that the volume of such region may correspond to a molecule of at least about 600 amu and often up to about 1600 amu or more.

Alternatively, a druggable region may be characterized by comparison to other regions on the same or other molecules. For example, the term "affinity region" refers to a druggable region on a molecule (such as a polypeptide of the invention) that is present in several other molecules, in so much as the structures of the same affinity regions are sufficiently the same so that they are expected to bind the same or related structural analogs. An example of an affinity region is an ATP-binding site of a protein kinase that is found in several protein kinases (whether or not of the same origin). The term "selectivity region" refers to a druggable region of a molecule that may not be found on other

molecules, in so much as the structures of different selectivity regions are sufficiently different so that they are not expected to bind the same or related structural analogs. An exemplary selectivity region is a catalytic domain of a protein kinase that exhibits specificity for one substrate. In certain instances, a single modulator may bind to the same affinity region across a number of proteins that have a substantially similar biological function, whereas the same modulator may bind to only one selectivity region of one of those proteins.

Continuing with examples of different druggable regions, the term "undesired region" refers to a druggable region of a molecule that upon interacting with another molecule results in an undesirable affect. For example, a binding site that oxidizes the interacting molecule (such as P-450 activity) and thereby results in increased toxicity for the oxidized molecule may be deemed a "undesired region". Other examples of potential undesired regions includes regions that upon interaction with a drug decrease the membrane permeability of the drug, increase the excretion of the drug, or increase the blood brain transport of the drug. It may be the case that, in certain circumstances, an undesired region will no longer be deemed an undesired region because the affect of the region will be favorable, e.g., a drug intended to treat a brain condition would benefit from interacting with a region that resulted in increased blood brain transport, whereas the same region could be deemed undesirable for drugs that were not intended to be delivered to the brain.

When used in reference to a druggable region, the "selectivity" or "specificity' of a molecule such as a modulator to a druggable region may be used to describe the binding between the molecule and a druggable region. For example, the selectivity of a modulator with respect to a druggable region may be expressed by comparison to another modulator, using the respective values of Kd (i.e., the dissociation constants for each modulator-druggable region complex) or, in cases where a biological effect is observed below the Kd, the ratio of the respective EC50's (i.e., the concentrations that produce 50% of the maximum response for the modulator interacting with each druggable region).

A "fusion protein" or "fusion polypeptide" refers to a chimeric protein as that term is known in the art and may be constructed using methods known in the art. In many examples of fusion proteins, there are two different polypeptide sequences, and in certain cases, there may be more. The sequences may be linked in frame. A fusion protein may include a domain which is found (albeit in a different protein) in an organism which also expresses the first protein, or it may be an "interspecies", "intergenic", etc. fusion expressed

by different kinds of organisms. In various embodiments, the fusion polypeptide may comprise one or more amino acid sequences linked to a first polypeptide. In the case where more than one amino acid sequence is fused to a first polypeptide, the fusion sequences may be multiple copies of the same sequence, or alternatively, may be different amino acid sequences. The fusion polypeptides may be fused to the N-terminus, the C-terminus, or the N- and C-terminus of the first polypeptide. Exemplary fusion proteins include polypeptides comprising a glutathione S-transferase tag (GST-tag), histidine tag (His-tag), an immunoglobulin domain or an immunoglobulin binding domain.

The term "gene" refers to a nucleic acid comprising an open reading frame encoding a polypeptide having exon sequences and optionally intron sequences. The term "intron" refers to a DNA sequence present in a given gene which is not translated into protein and is generally found between exons.

The term "having substantially similar biological activity", when used in reference to two polypeptides, refers to a biological activity of a first polypeptide which is substantially similar to at least one of the biological activities of a second polypeptide. A substantially similar biological activity means that the polypeptides carry out a similar function, e.g., a similar enzymatic reaction or a similar physiological process, etc. For example, two homologous proteins may have a substantially similar biological activity if they are involved in a similar enzymatic reaction, e.g., they are both kinases which catalyze phosphorylation of a substrate polypeptide, however, they may phosphorylate different regions on the same protein substrate or different substrate proteins altogether. Alternatively, two homologous proteins may also have a substantially similar biological activity if they are both involved in a similar physiological process, e.g., transcription. For example, two proteins may be transcription factors, however, they may bind to different DNA sequences or bind to different polypeptide interactors. Substantially similar biological activities may also be associated with proteins carrying out a similar structural role, for example, two membrane proteins.

The term "isolated polypeptide" refers to a polypeptide, in certain embodiments prepared from recombinant DNA or RNA, or of synthetic origin, or some combination thereof, which (1) is not associated with proteins that it is normally found with in nature, (2) is isolated from the cell in which it normally occurs, (3) is isolated free of other proteins from the same cellular source, e.g. free of other *P. aeruginosa* proteins, (4) is expressed by a cell from a different species, or (5) does not occur in nature.

The term "isolated nucleic acid" refers to a polynucleotide of genomic, cDNA, or synthetic origin or some combination there of, which (1) is not associated with the cell in which the "isolated nucleic acid" is found in nature, or (2) is operably linked to a polynucleotide to which it is not linked in nature.

The terms "label" or "labeled" refer to incorporation or attachment, optionally covalently or non-covalently, of a detectable marker into a molecule, such as a polypeptide. Various methods of labeling polypeptides are known in the art and may be used. Examples of labels for polypeptides include, but are not limited to, the following: radioisotopes, fluorescent labels, heavy atoms, enzymatic labels or reporter genes, chemiluminescent groups, biotinyl groups, predetermined polypeptide epitopes recognized by a secondary reporter (e.g., leucine zipper pair sequences, binding sites for secondary antibodies, metal binding domains, epitope tags). Examples and use of such labels are described in more detail below. In some embodiments, labels are attached by spacer arms of various lengths to reduce potential steric hindrance.

The term "mammal" is known in the art, and exemplary mammals include humans, primates, bovines, porcines, canines, felines, and rodents (e.g., mice and rats).

The term "modulation", when used in reference to a functional property or biological activity or process (e.g., enzyme activity or receptor binding), refers to the capacity to either up regulate (e.g., activate or stimulate), down regulate (e.g., inhibit or suppress) or otherwise change a quality of such property, activity or process. In certain instances, such regulation may be contingent on the occurrence of a specific event, such as activation of a signal transduction pathway, and/or may be manifest only in particular cell types.

The term "modulator" refers to a polypeptide, nucleic acid, macromolecule, complex, molecule, small molecule, compound, species or the like (naturally-occurring or non-naturally-occurring), or an extract made from biological materials such as bacteria, plants, fungi, or animal cells or tissues, that may be capable of causing modulation. Modulators may be evaluated for potential activity as inhibitors or activators (directly or indirectly) of a functional property, biological activity or process, or combination of them, (e.g., agonist, partial antagonist, partial agonist, inverse agonist, antagonist, anti-microbial agents, inhibitors of microbial infection or proliferation, and the like) by inclusion in assays. In such assays, many modulators may be screened at one time. The activity of a modulator may be known, unknown or partially known.

The term "motif" refers to an amino acid sequence that is commonly found in a protein of a particular structure or function. Typically, a consensus sequence is defined to represent a particular motif. The consensus sequence need not be strictly defined and may contain positions of variability, degeneracy, variability of length, etc. The consensus sequence may be used to search a database to identify other proteins that may have a similar structure or function due to the presence of the motif in its amino acid sequence. For example, on-line databases may be searched with a consensus sequence in order to identify other proteins containing a particular motif. Various search algorithms and/or programs may be used, including FASTA, BLAST or ENTREZ. FASTA and BLAST are available as a part of the GCG sequence analysis package (University of Wisconsin, Madison, Wis.). ENTREZ is available through the National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD.

The term "naturally-occurring", as applied to an object, refers to the fact that an object may be found in nature. For example, a polypeptide or polynucleotide sequence that is present in an organism (including bacteria) that may be isolated from a source in nature and which has not been intentionally modified by man in the laboratory is naturally-occurring.

The term "nucleic acid" refers to a polymeric form of nucleotides, either ribonucleotides or deoxynucleotides or a modified form of either type of nucleotide. The terms should also be understood to include, as equivalents, analogs of either RNA or DNA made from nucleotide analogs, and, as applicable to the embodiment being described, single-stranded (such as sense or antisense) and double-stranded polynucleotides.

The term "nucleic acid of the invention" refers to a nucleic acid encoding a polypeptide of the invention, e.g., a nucleic acid comprising a sequence consisting of, or consisting essentially of, the polynucleotide sequence set forth in SEQ ID NO: 1 or SEQ ID NO: 3. A nucleic acid of the invention may comprise all, or a portion of: the nucleotide sequence of SEQ ID NO: 1 or SEQ ID NO: 3; a nucleotide sequence at least 60%, 70%, 80%, 90%, 95%, 96%, 97%, 98% or 99% identical to SEQ ID NO: 1 or SEQ ID NO: 3; a nucleotide sequence that hybridizes under stringent conditions to SEQ ID NO: 1 or SEQ ID NO: 3; nucleotide sequences encoding polypeptides that are functionally equivalent to polypeptides of the invention; nucleotide sequences encoding polypeptides at least about 60%, 70%, 80%, 85%, 90%, 95%, 98%, 99% homologous or identical with an amino acid

sequence of SEQ ID NO: 2 or SEQ ID NO: 4; nucleotide sequences encoding polypeptides having an activity of a polypeptide of the invention and having at least about 60%, 70%, 80%, 85%, 90%, 95%, 98%, 99% or more homology or identity with SEQ ID NO: 2 or SEQ ID NO: 4; nucleotide sequences that differ by 1 to about 2, 3, 5, 7, 10, 15, 20, 30, 50, 75 or more nucleotide substitutions, additions or deletions, such as allelic variants, of SEQ ID NO: 1 and SEQ ID NO: 3; nucleic acids derived from and evolutionarily related to SEQ ID NO: 1 or SEQ ID NO: 3; and complements of, and nucleotide sequences resulting from the degeneracy of the genetic code, for all of the foregoing and other nucleic acids of the invention. Nucleic acids of the invention also include homologs, e.g., orthologs and paralogs, of SEQ ID NO: 1 or SEQ ID NO: 3 and also variants of SEQ ID NO: 1 or SEQ ID NO: 3 which have been codon optimized for expression in a particular organism (e.g., host cell).

The term "operably linked", when describing the relationship between two nucleic acid regions, refers to a juxtaposition wherein the regions are in a relationship permitting them to function in their intended manner. For example, a control sequence "operably linked" to a coding sequence is ligated in such a way that expression of the coding sequence is achieved under conditions compatible with the control sequences, such as when the appropriate molecules (e.g., inducers and polymerases) are bound to the control or regulatory sequence(s).

The term "phenotype" refers to the entire physical, biochemical, and physiological makeup of a cell, e.g., having any one trait or any group of traits.

The term "polypeptide", and the terms "protein" and "peptide" which are used interchangeably herein, refers to a polymer of amino acids. Exemplary polypeptides include gene products, naturally-occurring proteins, homologs, orthologs, paralogs, fragments, and other equivalents, variants and analogs of the foregoing.

The terms "polypeptide fragment" or "fragment", when used in reference to a reference polypeptide, refers to a polypeptide in which amino acid residues are deleted as compared to the reference polypeptide itself, but where the remaining amino acid sequence is usually identical to the corresponding positions in the reference polypeptide. Such deletions may occur at the amino-terminus or carboxy-terminus of the reference polypeptide, or alternatively both. Fragments typically are at least 5, 6, 8 or 10 amino acids long, at least 14 amino acids long, at least 20, 30, 40 or 50 amino acids long, at least 75 amino acids long, or at least 100, 150, 200, 300, 500 or more amino acids long. A fragment

can retain one or more of the biological activities of the reference polypeptide. In certain embodiments, a fragment may comprise a druggable region, and optionally additional amino acids on one or both sides of the druggable region, which additional amino acids may number from 5, 10, 15, 20, 30, 40, 50, or up to 100 or more residues. Further, fragments can include a sub-fragment of a specific region, which sub-fragment retains a function of the region from which it is derived. In another embodiment, a fragment may have immunogenic properties.

The term "polypeptide of the invention" refers to a polypeptide comprising the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4, or an equivalent or fragment thereof, e.g., a polypeptide comprising a sequence consisting of, or consisting essentially of, the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4. Polypeptides of the invention include polypeptides comprising all or a portion of the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4 with 1 to about 2, 3, 5, 7, 10, 15, 20, 30, 50, 75 or more conservative amino acid substitutions; an amino acid sequence that is at least 60%, 70%, 80%, 90%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO: 2 or SEQ ID NO: 4; and functional fragments thereof. Polypeptides of the invention also include homologs, e.g., orthologs and paralogs, of SEQ ID NO: 2 or SEQ ID NO: 4.

The term "purified" refers to an object species that is the predominant species present (i.e., on a molar basis it is more abundant than any other individual species in the composition). A "purified fraction" is a composition wherein the object species comprises at least about 50 percent (on a molar basis) of all species present. In making the determination of the purity of a species in solution or dispersion, the solvent or matrix in which the species is dissolved or dispersed is usually not included in such determination; instead, only the species (including the one of interest) dissolved or dispersed are taken into account. Generally, a purified composition will have one species that comprises more than about 80 percent of all species present in the composition, more than about 85%, 90%, 95%, 99% or more of all species present. The object species may be purified to essential homogeneity (contaminant species cannot be detected in the composition by conventional detection methods) wherein the composition consists essentially of a single species. A skilled artisan may purify a polypeptide of the invention using standard techniques for protein purification in light of the teachings herein. Purity of a polypeptide may be

determined by a number of methods known to those of skill in the art, including for example, amino-terminal amino acid sequence analysis, gel electrophoresis, mass-spectrometry analysis and the methods described in the Exemplification section herein.

The terms "recombinant protein" or "recombinant polypeptide" refer to a polypeptide which is produced by recombinant DNA techniques. An example of such techniques includes the case when DNA encoding the expressed protein is inserted into a suitable expression vector which is in turn used to transform a host cell to produce the protein or polypeptide encoded by the DNA.

A "reference sequence" is a defined sequence used as a basis for a sequence comparison; a reference sequence may be a subset of a larger sequence, for example, as a segment of a full-length protein given in a sequence listing such as SEQ ID NO: 2 or SEQ ID NO: 4, or may comprise a complete protein sequence. Generally, a reference sequence is at least 200, 300 or 400 nucleotides in length, frequently at least 600 nucleotides in length, and often at least 800 nucleotides in length (or the protein equivalent if it is shorter or longer in length). Because two proteins may each (1) comprise a sequence (i.e., a portion of the complete protein sequence) that is similar between the two proteins, and (2) may further comprise a sequence that is divergent between the two proteins, sequence comparisons between two (or more) proteins are typically performed by comparing sequences of the two proteins over a "comparison window" to identify and compare local regions of sequence similarity.

The term "regulatory sequence" is a generic term used throughout the specification to refer to polynucleotide sequences, such as initiation signals, enhancers, regulators and promoters, that are necessary or desirable to affect the expression of coding and non-coding sequences to which they are operably linked. Exemplary regulatory sequences are described in Goeddel; *Gene Expression Technology*: *Methods in Enzymology*, Academic Press, San Diego, CA (1990), and include, for example, the early and late promoters of SV40, adenovirus or cytomegalovirus immediate early promoter, the lac system, the trp system, the TAC or TRC system, T7 promoter whose expression is directed by T7 RNA polymerase, the major operator and promoter regions of phage lambda, the control regions for fd coat protein, the promoter for 3-phosphoglycerate kinase or other glycolytic enzymes, the promoters of acid phosphatase, e.g., Pho5, the promoters of the yeast α-mating factors, the polyhedron promoter of the baculovirus system and other sequences known to control

the expression of genes of prokaryotic or eukaryotic cells or their viruses, and various combinations thereof. The nature and use of such control sequences may differ depending upon the host organism. In prokaryotes, such regulatory sequences generally include promoter, ribosomal binding site, and transcription termination sequences. The term "regulatory sequence" is intended to include, at a minimum, components whose presence may influence expression, and may also include additional components whose presence is advantageous, for example, leader sequences and fusion partner sequences. In certain embodiments, transcription of a polynucleotide sequence is under the control of a promoter sequence (or other regulatory sequence) which controls the expression of the polynucleotide in a cell-type in which expression is intended. It will also be understood that the polynucleotide can be under the control of regulatory sequences which are the same or different from those sequences which control expression of the naturally-occurring form of the polynucleotide.

The term "reporter gene" refers to a nucleic acid comprising a nucleotide sequence encoding a protein that is readily detectable either by its presence or activity, including, but not limited to, luciferase, fluorescent protein (e.g., green fluorescent protein), chloramphenicol acetyl transferase, β-galactosidase, secreted placental alkaline phosphatase, β-lactamase, human growth hormone, and other secreted enzyme reporters. Generally, a reporter gene encodes a polypeptide not otherwise produced by the host cell, which is detectable by analysis of the cell(s), e.g., by the direct fluorometric, radioisotopic or spectrophotometric analysis of the cell(s) and preferably without the need to kill the cells for signal analysis. In certain instances, a reporter gene encodes an enzyme, which produces a change in fluorometric properties of the host cell, which is detectable by qualitative, quantitative or semiquantitative function or transcriptional activation. Exemplary enzymes include esterases, β-lactamase, phosphatases, peroxidases, proteases (tissue plasminogen activator or urokinase) and other enzymes whose function may be detected by appropriate chromogenic or fluorogenic substrates known to those skilled in the art or developed in the future.

The term "sequence homology" refers to the proportion of base matches between two nucleic acid sequences or the proportion of amino acid matches between two amino acid sequences. When sequence homology is expressed as a percentage, e.g., 50%, the percentage denotes the proportion of matches over the length of sequence from a desired

sequence (e.g., SEQ. ID NO: 1) that is compared to some other sequence. Gaps (in either of the two sequences) are permitted to maximize matching; gap lengths of 15 bases or less are usually used, 6 bases or less are used more frequently, with 2 bases or less used even more frequently. The term "sequence identity" means that sequences are identical (i.e., on a nucleotide-by-nucleotide basis for nucleic acids or amino acid-by-amino acid basis for polypeptides) over a window of comparison. The term "percentage of sequence identity" is calculated by comparing two optimally aligned sequences over the comparison window, determining the number of positions at which the identical amino acids occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the comparison window, and multiplying the result by 100 to yield the percentage of sequence identity. Methods to calculate sequence identity are known to those of skill in the art and described in further detail below.

The term "small molecule" refers to a compound, which has a molecular weight of less than about 5 kD, less than about 2.5 kD, less than about 1.5 kD, or less than about 0.9 kD. Small molecules may be, for example, nucleic acids, peptides, polypeptides, peptide nucleic acids, peptidomimetics, carbohydrates, lipids or other organic (carbon containing) or inorganic molecules. Many pharmaceutical companies have extensive libraries of chemical and/or biological mixtures, often fungal, bacterial, or algal extracts, which can be screened with any of the assays of the invention. The term "small organic molecule" refers to a small molecule that is often identified as being an organic or medicinal compound, and does not include molecules that are exclusively nucleic acids, peptides or polypeptides.

The term "soluble" as used herein with reference to a polypeptide of the invention or other protein, means that upon expression in cell culture, at least some portion of the polypeptide or protein expressed remains in the cytoplasmic fraction of the cell and does not fractionate with the cellular debris upon lysis and centrifugation of the lysate. Solubility of a polypeptide may be increased by a variety of art recognized methods, including fusion to a heterologous amino acid sequence, deletion of amino acid residues, amino acid substitution (e.g., enriching the sequence with amino acid residues having hydrophilic side chains), and chemical modification (e.g., addition of hydrophilic groups). The solubility of polypeptides may be measured using a variety of art recognized techniques, including, dynamic light scattering to determine aggregation state, UV absorption, centrifugation to separate aggregated from non-aggregated material, and SDS gel electrophoresis (e.g., the amount of protein in the soluble fraction is compared to the

amount of protein in the soluble and insoluble fractions combined). When expressed in a host cell, the polypeptides of the invention may be at least about 1%, 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% or more soluble, e.g., at least about 1%, 2%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% or more of the total amount of protein expressed in the cell is found in the cytoplasmic fraction. In certain embodiments, a one liter culture of cells expressing a polypeptide of the invention will produce at least about 0.1, 0.2, 0.5, 1, 2, 5, 10, 20, 30, 40, 50 milligrams or more of soluble protein. In an exemplary embodiment, a polypeptide of the invention is at least about 10% soluble and will produce at least about 1 milligram of protein from a one liter cell culture.

The term "specifically hybridizes" refers to detectable and specific nucleic acid binding. Polynucleotides, oligonucleotides and nucleic acids of the invention selectively hybridize to nucleic acid strands under hybridization and wash conditions that minimize appreciable amounts of detectable binding to nonspecific nucleic acids. Stringent conditions may be used to achieve selective hybridization conditions as known in the art and discussed herein. Generally, the nucleic acid sequence homology between the polynucleotides, oligonucleotides, and nucleic acids of the invention and a nucleic acid sequence of interest will be at least 30%, 40%, 50%, 60%, 70%, 80%, 85%, 90%, 95%, 98%, 99%, or more. In certain instances, hybridization and washing conditions are performed under stringent conditions according to conventional hybridization procedures and as described further herein.

The terms "stringent conditions" or "stringent hybridization conditions" refer to conditions which promote specific hydribization between two complementary polynucleotide strands so as to form a duplex. Stringent conditions may be selected to be about 5°C lower than the thermal melting point (Tm) for a given polynucleotide duplex at a defined ionic strength and pH. The length of the complementary polynucleotide strands and their GC content will determine the Tm of the duplex, and thus the hybridization conditions necessary for obtaining a desired specificity of hybridization. The Tm is the temperature (under defined ionic strength and pH) at which 50% of the a polynucleotide sequence hybridizes to a perfectly matched complementary strand. In certain cases it may be desirable to increase the stringency of the hybridization conditions to be about equal to the Tm for a particular duplex.

A variety of techniques for estimating the Tm are available. Typically, G-C base pairs in a duplex are estimated to contribute about 3°C to the Tm, while A-T base pairs are

estimated to contribute about 2°C, up to a theoretical maximum of about 80-100°C. However, more sophisticated models of Tm are available in which G-C stacking interactions, solvent effects, the desired assay temperature and the like are taken into account. For example, probes can be designed to have a dissociation temperature (Td) of approximately 60°C, using the formula: Td = (((((3 x #GC) + (2 x #AT)) x 37) - 562)/#bp) - 5; where #GC, #AT, and #bp are the number of guanine-cytosine base pairs, the number of adenine-thymine base pairs, and the number of total base pairs, respectively, involved in the formation of the duplex.

Hybridization may be carried out in 5xSSC, 4xSSC, 3xSSC, 2xSSC, 1xSSC or 0.2xSSC for at least about 1 hour, 2 hours, 5 hours, 12 hours, or 24 hours. The temperature of the hybridization may be increased to adjust the stringency of the reaction, for example, from about 25°C (room temperature), to about 45°C, 50°C, 55°C, 60°C, or 65°C. The hybridization reaction may also include another agent affecting the stringency, for example, hybridization conducted in the presence of 50% formamide increases the stringency of hybridization at a defined temperature.

The hybridization reaction may be followed by a single wash step, or two or more wash steps, which may be at the same or a different salinity and temperature. For example, the temperature of the wash may be increased to adjust the stringency from about 25°C (room temperature), to about 45°C, 50°C, 55°C, 60°C, 65°C, or higher. The wash step may be conducted in the presence of a detergent, e.g., 0.1 or 0.2% SDS. For example, hybridization may be followed by two wash steps at 65°C each for about 20 minutes in 2xSSC, 0.1% SDS, and optionally two additional wash steps at 65°C each for about 20 minutes in 0.2xSSC, 0.1%SDS.

Exemplary stringent hybridization conditions include overnight hybridization at 65°C in a solution comprising, or consisting of, 50% formamide, 10xDenhardt (0.2% Ficoll, 0.2% Polyvinylpyrrolidone, 0.2% bovine serum albumin) and 200 g/ml of denatured carrier DNA, e.g., sheared salmon sperm DNA, followed by two wash steps at 65°C each for about 20 minutes in 2xSSC, 0.1% SDS, and two wash steps at 65°C each for about 20 minutes in 0.2xSSC, 0.1%SDS.

Hybridization may consist of hybridizing two nucleic acids in solution, or a nucleic acid in solution to a nucleic acid attached to a solid support, e.g., a filter. When one nucleic acid is on a solid support, a prehybridization step may be conducted prior to hybridization. Prehybridization may be carried out for at least about 1 hour, 3 hours or 10 hours in the

same solution and at the same temperature as the hybridization solution (without the complementary polynucleotide strand).

Appropriate stringency conditions are known to those skilled in the art or may be determined experimentally by the skilled artisan. See, for example, Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1-12.3.6; Sambrook et al., 1989, Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, N.Y; S. Agrawal (ed.) Methods in Molecular Biology, volume 20; Tijssen (1993) Laboratory Techniques in biochemistry and molecular biology-hybridization with nucleic acid probes, e.g., part I chapter 2 "Overview of principles of hybridization and the strategy of nucleic acid probe assays", Elsevier, New York; and Tibanyenda, N. et al., Eur. J. Biochem. 139:19 (1984) and Ebel, S. et al., Biochem. 31:12083 (1992).

As applied to proteins, the term "substantial identity" means that two protein sequences, when optimally aligned, such as by the programs GAP or BESTFIT using default gap weights, typically share at least about 70 percent sequence identity, alternatively at least about 80, 85, 90, 95 percent sequence identity or more. In certain instances, residue positions that are not identical differ by conservative amino acid substitutions, which are described above.

The term "structural motif", when used in reference to a polypeptide, refers to a polypeptide that, although it may have different amino acid sequences, may result in a similar structure, wherein by structure is meant that the motif forms generally the same tertiary structure, or that certain amino acid residues within the motif, or alternatively their backbone or side chains (which may or may not include the C atoms of the side chains) are positioned in a like relationship with respect to one another in the motif.

The term "test compound" refers to a molecule to be tested by one or more screening method(s) as a putative modulator of a polypeptide of the invention or other biological entity or process. A test compound is usually not known to bind to a target of interest. The term "control test compound" refers to a compound known to bind to the target (e.g., a known agonist, antagonist, partial agonist or inverse agonist). The term "test compound" does not include a chemical added as a control condition that alters the function of the target to determine signal specificity in an assay. Such control chemicals or conditions include chemicals that 1) nonspecifically or substantially disrupt protein structure (e.g., denaturing agents (e.g., urea or guanidinium), chaotropic agents, sulfhydryl reagents (e.g., dithiothreitol and β-mercaptoethanol), and proteases), 2) generally inhibit

cell metabolism (e.g., mitochondrial uncouplers) and 3) non-specifically disrupt electrostatic or hydrophobic interactions of a protein (e.g., high salt concentrations, or detergents at concentrations sufficient to non-specifically disrupt hydrophobic interactions). Further, the term "test compound" also does not include compounds known to be unsuitable for a therapeutic use for a particular indication due to toxicity of the subject. In certain embodiments, various predetermined concentrations of test compounds are used for screening such as 0.01 μM, 0.1 μM, 1.0 μM, and 10.0 μM. Examples of test compounds include, but are not limited to, peptides, nucleic acids, carbohydrates, and small molecules. The term "novel test compound" refers to a test compound that is not in existence as of the filing date of this application. In certain assays using novel test compounds, the novel test compounds comprise at least about 50%, 75%, 85%, 90%, 95% or more of the test compounds used in the assay or in any particular trial of the assay.

The term "therapeutically effective amount" refers to that amount of a modulator, drug or other molecule which is sufficient to effect treatment when administered to a subject in need of such treatment. The therapeutically effective amount will vary depending upon the subject and disease condition being treated, the weight and age of the subject, the severity of the disease condition, the manner of administration and the like, which can readily be determined by one of ordinary skill in the art.

The term "transfection" means the introduction of a nucleic acid, e.g., an expression vector, into a recipient cell, which in certain instances involves nucleic acid-mediated gene transfer. The term "transformation" refers to a process in which a cell's genotype is changed as a result of the cellular uptake of exogenous nucleic acid. For example, a transformed cell may express a recombinant form of a polypeptide of the invention or antisense expression may occur from the transferred gene so that the expression of a naturally-occurring form of the gene is disrupted.

The term "transgene" means a nucleic acid sequence, which is partly or entirely heterologous to a transgenic animal or cell into which it is introduced, or, is homologous to an endogenous gene of the transgenic animal or cell into which it is introduced, but which is designed to be inserted, or is inserted, into the animal's genome in such a way as to alter the genome of the cell into which it is inserted (e.g., it is inserted at a location which differs from that of the natural gene or its insertion results in a knockout). A transgene may include one or more regulatory sequences and any other nucleic acids, such as introns, that may be necessary for optimal expression.

The term "transgenic animal" refers to any animal, for example, a mouse, rat or other non-human mammal, a bird or an amphibian, in which one or more of the cells of the animal contain heterologous nucleic acid introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the cell, directly or indirectly, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. The term genetic manipulation does not include classical cross-breeding, or *in vitro* fertilization, but rather is directed to the introduction of a recombinant DNA molecule. This molecule may be integrated within a chromosome, or it may be extrachromosomally replicating DNA. In the typical transgenic animals described herein, the transgene causes cells to express a recombinant form of a protein. However, transgenic animals in which the recombinant gene is silent are also contemplated.

The term "vector" refers to a nucleic acid capable of transporting another nucleic acid to which it has been linked. One type of vector which may be used in accord with the invention is an episome, i.e., a nucleic acid capable of extra-chromosomal replication. Other vectors include those capable of autonomous replication and expression of nucleic acids to which they are linked. Vectors capable of directing the expression of genes to which they are operatively linked are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of "plasmids" which refer to circular double stranded DNA molecules which, in their vector form are not bound to the chromosome. In the present specification, "plasmid" and "vector" are used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors which serve equivalent functions and which become known in the art subsequently hereto.

Unless otherwise indicated, all numbers expressing quantities of ingredients, reaction conditions, and so forth used in the specification and claims are to be understood as being modified in all instances by the term "about." Accordingly, unless indicated to the contrary, the numerical parameters set forth in this specification and attached claims are approximations that may vary depending upon the desired properties sought to be obtained by the present invention.

2. Polypeptides of the Invention

The present invention makes available in a variety of embodiments soluble, purified and/or isolated forms of the polypeptides of the invention. Milligram quantities of an

exemplary polypeptide of the invention, SEQ ID NO: 4 (optionally with a tag, and optionally labeled), have been isolated in a highly purified form. The present invention provides for expressing and purifying polypeptides of the invention in quantities that equal or exceed the quantity of polypeptide(s) of the invention expressed and purified as provided in the Exemplification section below (or smaller amount(s) thereof, such as 25%, 33%, 50% or 75% of the amount(s) so expressed and/or purified).

In one aspect, the present invention contemplates an isolated polypeptide comprising (a) the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4, (b) the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4 with 1 to about 20 conservative amino acid substitutions, deletions or additions, (c) an amino acid sequence that is at least 90% identical to SEQ ID NO: 2 or SEQ ID NO: 4 or (d) a functional fragment of a polypeptide having an amino acid sequence set forth in (a), (b) or (c). In another aspect, the present invention contemplates a composition comprising such an isolated polypeptide and less than about 10%, or alternatively 5%, or alternatively 1%, contaminating biological macromolecules or polypeptides.

It may be the case that the amino acid sequence of SEQ ID NO: 4 differs from that of SEQ ID NO: 2 by one or more amino acids. SEQ ID NO: 4 is determined from the experimentally determined nucleic acid sequence SED ID NO: 3, and SEQ ID NO: 2 is determined from SEQ ID NO: 1, which is obtained as described in EXAMPLE 1. In such a case, the present invention contemplates the specific amino acid sequences of SEQ ID NO: 2 and SEQ ID NO: 4, and variants thereof, as well as any differences (if any) in the polypeptides of the invention based on those SEQ ID NOS and nucleic acid sequences encoding the same.

In certain embodiments, a polypeptide of the invention is a fusion protein containing a domain which increases its solubility and/or facilitates its purification, identification, detection, and/or structural characterization. Exemplary domains, include, for example, glutathione S-transferase (GST), protein A, protein G, calmodulin-binding peptide, thioredoxin, maltose binding protein, HA, myc, poly arginine, poly His, poly His-Asp or FLAG fusion proteins and tags. Additional exemplary domains include domains that alter protein localization *in vivo*, such as signal peptides, type III secretion system-targeting peptides, transcytosis domains, nuclear localization signals, etc. In various embodiments, a polypeptide of the invention may comprise one or more heterologous fusions. Polypeptides may contain multiple copies of the same fusion domain or may contain fusions to two or

more different domains. The fusions may occur at the N-terminus of the polypeptide, at the C-terminus of the polypeptide, or at both the N- and C-terminus of the polypeptide. It is also within the scope of the invention to include linker sequences between a polypeptide of the invention and the fusion domain in order to facilitate construction of the fusion protein or to optimize protein expression or structural constraints of the fusion protein. In another embodiment, the polypeptide may be constructed so as to contain protease cleavage sites between the fusion polypeptide and polypeptide of the invention in order to remove the tag after protein expression or thereafter. Examples of suitable endoproteases, include, for example, Factor Xa and TEV proteases.

In another embodiment, a polypeptide of the invention may be modified so that its rate of traversing the cellular membrane is increased. For example, the polypeptide may be fused to a second peptide which promotes "transcytosis," e.g., uptake of the peptide by cells. The peptide may be a portion of the HIV transactivator (TAT) protein, such as the fragment corresponding to residues 37-62 or 48-60 of TAT, portions which have been observed to be rapidly taken up by a cell *in vitro* (Green and Loewenstein, (1989) Cell 55:1179-1188). Alternatively, the internalizing peptide may be derived from the *Drosophila antennapedia* protein, or homologs thereof. The 60 amino acid long homeodomain of the homeo-protein antennapedia has been demonstrated to translocate through biological membranes and can facilitate the translocation of heterologous polypeptides to which it is coupled. Thus, polypeptides may be fused to a peptide consisting of about amino acids 42-58 of *Drosophila antennapedia* or shorter fragments for transcytosis (Derossi et al. (1996) J Biol Chem 271:18188-18193; Derossi et al. (1994) J Biol Chem 269:10444-10450; and Perez et al. (1992) J Cell Sci 102:717-722). The transcytosis polypeptide may also be a non-naturally-occurring membrane-translocating sequence (MTS), such as the peptide sequences disclosed in U.S. Patent No. 6,248,558.

In another embodiment, a polypeptide of the invention is labeled with an isotopic label to facilitate its detection and or structural characterization using nuclear magnetic resonance or another applicable technique. Exemplary isotopic labels include radioisotopic labels such as, for example, potassium-40 (^{40}K), carbon-14 (^{14}C), tritium (^{3}H), sulphur-35 (^{35}S), phosphorus-32 (^{32}P), technetium-99m (^{99m}Tc), thallium-201 (^{201}Tl), gallium-67 (^{67}Ga), indium-111 (^{111}In), iodine-123 (^{123}I), iodine-131 (^{131}I), yttrium-90 (^{90}Y), samarium-153 (^{153}Sm), rhenium-186 (^{186}Re), rhenium-188 (^{188}Re), dysprosium-165 (^{165}Dy) and holmium-166 (^{166}Ho). The isotopic label may also be an atom with non zero nuclear spin,

including, for example, hydrogen-1 (^{1}H), hydrogen-2 (^{2}H), hydrogen-3 (^{3}H), phosphorous-31 (^{31}P), sodium-23 (^{23}Na), nitrogen-14 (^{14}N), nitrogen-15 (^{15}N), carbon-13 (^{13}C) and fluorine-19 (^{19}F). In certain embodiments, the polypeptide is uniformly labeled with an isotopic label, for example, wherein at least 50%, 70%, 80%, 90%, 95%, or 98% of the possible labels in the polypeptide are labeled, e.g., wherein at least 50%, 70%, 80%, 90%, 95%, or 98% of the nitrogen atoms in the polypeptide are ^{15}N, and/or wherein at least 50%, 70%, 80%, 90%, 95%, or 98% of the carbon atoms in the polypeptide are ^{13}C, and/or wherein at least 50%, 70%, 80%, 90%, 95%, or 98% of the hydrogen atoms in the polypeptide are ^{2}H. In other embodiments, the isotopic label is located in one or more specific locations within the polypeptide, for example, the label may be specifically incorporated into one or more of the leucine residues of the polypeptide. The invention also encompasses the embodiment wherein a single polypeptide comprises two, three or more different isotopic labels, for example, the polypeptide comprises both ^{15}N and ^{13}C labeling.

In yet another embodiment, the polypeptides of the invention are labeled to facilitate structural characterization using x-ray crystallography or another applicable technique. Exemplary labels include heavy atom labels such as, for example, cobalt, selenium, krypton, bromine, strontium, molybdenum, ruthenium, rhodium, palladium, silver, cadmium, tin, iodine, xenon, barium, lanthanum, cerium, praseodymium, neodymium, samarium, europium, gadolinium, terbium, dysprosium, holmium, erbium, thulium, ytterbium, lutetium, tantalum, tungsten, rhenium, osmium, iridium, platinum, gold, mercury, thallium, lead, thorium and uranium. In an exemplary embodiment, the polypeptide is labeled with seleno-methionine.

A variety of methods are available for preparing a polypeptide with a label, such as a radioisotopic label or heavy atom label. For example, in one such method, an expression vector comprising a nucleic acid encoding a polypeptide is introduced into a host cell, and the host cell is cultured in a cell culture medium in the presence of a source of the label, thereby generating a labeled polypeptide. As indicated above, the extent to which a polypeptide may be labeled may vary.

In still another embodiment, the polypeptides of the invention are labeled with a fluorescent label to facilitate their detection, purification, or structural characterization. In an exemplary embodiment, a polypeptide of the invention is fused to a heterologous polypeptide sequence which produces a detectable fluorescent signal, including, for example, green fluorescent protein (GFP), enhanced green fluorescent protein (EGFP),

Renilla Reniformis green fluorescent protein, GFPmut2, GFPuv4, enhanced yellow fluorescent protein (EYFP), enhanced cyan fluorescent protein (ECFP), enhanced blue fluorescent protein (EBFP), citrine and red fluorescent protein from discosoma (dsRED).

In other embodiments, the invention provides for polypeptides of the invention immobilized onto a solid surface, including, microtiter plates, slides, beads, films, etc. The polypeptides of the invention may be immobilized onto a "chip" as part of an array. An array, having a plurality of addresses, may comprise one or more polypeptides of the invention in one or more of those addresses. In one embodiment, the chip comprises one or more polypeptides of the invention as part of an array of *P. aeruginosa* polypeptide sequences.

In other embodiments, the invention provides for polypeptides of the invention immobilized onto a solid surface, including, plates, microtiter plates, slides, beads, particles, spheres, films, strands, precipitates, gels, sheets, tubing, containers, capillaries, pads, slices, etc. The polypeptides of the invention may be immobilized onto a "chip" as part of an array. An array, having a plurality of addresses, may comprise one or more polypeptides of the invention in one or more of those addresses. In one embodiment, the chip comprises one or more polypeptides of the invention as part of an array that contains at least some polypeptide sequences from *P. aeruginosa*.

In still other embodiments, the invention comprises the polypeptide sequences of the invention in computer readable format. The invention also encompasses a database comprising the polypeptide sequences of the invention.

In other embodiments, the invention relates to the polypeptides of the invention contained within a vessels useful for manipulation of the polypeptide sample. For example, the polypeptides of the invention may be contained within a microtiter plate to facilitate detection, screening or purification of the polypeptide. The polypeptides may also be contained within a syringe as a container suitable for administering the polypeptide to a subject in order to generate antibodies or as part of a vaccination regimen. The polypeptides may also be contained within an NMR tube in order to enable characterization by nuclear magnetic resonance techniques.

In still other embodiments, the invention relates to a crystallized polypeptide of the invention and crystallized polypeptides which have been mounted for examination by x-ray crystallography as described further below. In certain instances, a polypeptide of the invention in crystal form may be single crystals of various dimensions (e.g., micro-crystals)

or may be an aggregate of crystalline material. In another aspect, the present invention contemplates a crystallized complex including a polypeptide of the invention and one or more of the following: a co-factor (such as a salt, metal, nucleotide, oligonucleotide or polypeptide), a modulator, or a small molecule. In another aspect, the present invention contemplates a crystallized complex including a polypeptide of the invention and any other molecule or atom (such as a metal ion) that associates with the polypeptide *in vivo*.

In certain embodiments, polypeptides of the invention may be synthesized chemically, ribosomally in a cell free system, or ribosomally within a cell. Chemical synthesis of polypeptides of the invention may be carried out using a variety of art recognized methods, including stepwise solid phase synthesis, semi-synthesis through the conformationally-assisted re-ligation of peptide fragments, enzymatic ligation of cloned or synthetic peptide segments, and chemical ligation. Native chemical ligation employs a chemoselective reaction of two unprotected peptide segments to produce a transient thioester-linked intermediate. The transient thioester-linked intermediate then spontaneously undergoes a rearrangement to provide the full length ligation product having a native peptide bond at the ligation site. Full length ligation products are chemically identical to proteins produced by cell free synthesis. Full length ligation products may be refolded and/or oxidized, as allowed, to form native disulfide-containing protein molecules. (see e.g., U.S. Patent Nos. 6,184,344 and 6,174,530; and T. W. Muir et al., Curr. Opin. Biotech. (1993): vol. 4, p 420; M. Miller, et al., Science (1989): vol. 246, p 1149; A. Wlodawer, et al., Science (1989): vol. 245, p 616; L. H. Huang, et al., Biochemistry (1991): vol. 30, p 7402; M. Schnolzer, et al., Int. J. Pept. Prot. Res. (1992): vol. 40, p 180-193; K. Rajarathnam, et al., Science (1994): vol. 264, p 90; R. E. Offord, "Chemical Approaches to Protein Engineering", in Protein Design and the Development of New therapeutics and Vaccines, J. B. Hook, G. Poste, Eds., (Plenum Press, New York, 1990) pp. 253-282; C. J. A. Wallace, et al., J. Biol. Chem. (1992): vol. 267, p 3852; L. Abrahmsen, et al., Biochemistry (1991): vol. 30, p 4151; T. K. Chang, et al., Proc. Natl. Acad. Sci. USA (1994) 91: 12544-12548; M. Schnlzer, et al., Science (1992): vol., 3256, p 221; and K. Akaji, et al., Chem. Pharm. Bull. (Tokyo) (1985) 33: 184).

In certain embodiments, it may be advantageous to provide naturally-occurring or experimentally-derived homologs of a polypeptide of the invention. Such homologs may function in a limited capacity as a modulator to promote or inhibit a subset of the biological activities of the naturally-occurring form of the polypeptide. Thus, specific biological

effects may be elicited by treatment with a homolog of limited function, and with fewer side effects relative to treatment with agonists or antagonists which are directed to all of the biological activities of a polypeptide of the invention. For instance, antagonistic homologs may be generated which interfere with the ability of the wild-type polypeptide of the invention to associate with certain proteins, but which do not substantially interfere with the formation of complexes between the native polypeptide and other cellular proteins.

Another aspect of the invention relates to polypeptides derived from the full-length polypeptides of the invention. Isolated peptidyl portions of those polypeptides may be obtained by screening polypeptides recombinantly produced from the corresponding fragment of the nucleic acid encoding such polypeptides. In addition, fragments may be chemically synthesized using techniques known in the art such as conventional Merrifield solid phase f-Moc or t-Boc chemistry. For example, proteins may be arbitrarily divided into fragments of desired length with no overlap of the fragments, or may be divided into overlapping fragments of a desired length. The fragments may be produced (recombinantly or by chemical synthesis) and tested to identify those peptidyl fragments having a desired property, for example, the capability of functioning as a modulator of the polypeptides of the invention. In an illustrative embodiment, peptidyl portions of a protein of the invention may be tested for binding activity, as well as inhibitory ability, by expression as, for example, thioredoxin fusion proteins, each of which contains a discrete fragment of a protein of the invention (see, for example, U.S. Patents 5,270,181 and 5,292,646; and PCT publication WO94/ 02502).

In another embodiment, truncated polypeptides may be prepared. Truncated polypeptides have from 1 to 20 or more amino acid residues removed from either or both the N- and C-termini. Such truncated polypeptides may prove more amenable to expression, purification or characterization than the full-length polypeptide. For example, truncated polypeptides may prove more amenable than the full-length polypeptide to crystallization, to yielding high quality diffracting crystals or to yielding an HSQC spectrum with high intensity peaks and minimally overlapping peaks. In addition, the use of truncated polypeptides may also identify stable and active domains of the full-length polypeptide that may be more amenable to characterization.

It is also possible to modify the structure of the polypeptides of the invention for such purposes as enhancing therapeutic or prophylactic efficacy, or stability (e.g., *ex vivo* shelf life, resistance to proteolytic degradation *in vivo*, etc.). Such modified polypeptides,

when designed to retain at least one activity of the naturally-occurring form of the protein, are considered "functional equivalents" of the polypeptides described in more detail herein. Such modified polypeptides may be produced, for instance, by amino acid substitution, deletion, or addition, which substitutions may consist in whole or part by conservative amino acid substitutions.

For instance, it is reasonable to expect that an isolated conservative amino acid substitution, such as replacement of a leucine with an isoleucine or valine, an aspartate with a glutamate, a threonine with a serine, will not have a major affect on the biological activity of the resulting molecule. Whether a change in the amino acid sequence of a polypeptide results in a functional homolog may be readily determined by assessing the ability of the variant polypeptide to produce a response similar to that of the wild-type protein. Polypeptides in which more than one replacement has taken place may readily be tested in the same manner.

This invention further contemplates a method of generating sets of combinatorial mutants of polypeptides of the invention, as well as truncation mutants, and is especially useful for identifying potential variant sequences (e.g. homologs). The purpose of screening such combinatorial libraries is to generate, for example, homologs which may modulate the activity of a polypeptide of the invention, or alternatively, which possess novel activities altogether. Combinatorially-derived homologs may be generated which have a selective potency relative to a naturally-occurring protein. Such homologs may be used in the development of therapeutics.

Likewise, mutagenesis may give rise to homologs which have intracellular half-lives dramatically different than the corresponding wild-type protein. For example, the altered protein may be rendered either more stable or less stable to proteolytic degradation or other cellular process which result in destruction of, or otherwise inactivation of the protein. Such homologs, and the genes which encode them, may be utilized to alter protein expression by modulating the half-life of the protein. As above, such proteins may be used for the development of therapeutics or treatment.

In similar fashion, protein homologs may be generated by the present combinatorial approach to act as antagonists, in that they are able to interfere with the activity of the corresponding wild-type protein.

In a representative embodiment of this method, the amino acid sequences for a population of protein homologs are aligned, preferably to promote the highest homology

possible. Such a population of variants may include, for example, homologs from one or more species, or homologs from the same species but which differ due to mutation. Amino acids which appear at each position of the aligned sequences are selected to create a degenerate set of combinatorial sequences. In certain embodiments, the combinatorial library is produced by way of a degenerate library of genes encoding a library of polypeptides which each include at least a portion of potential protein sequences. For instance, a mixture of synthetic oligonucleotides may be enzymatically ligated into gene sequences such that the degenerate set of potential nucleotide sequences are expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (e.g. for phage display).

There are many ways by which the library of potential homologs may be generated from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence may be carried out in an automatic DNA synthesizer, and the synthetic genes may then be ligated into an appropriate vector for expression. One purpose of a degenerate set of genes is to provide, in one mixture, all of the sequences encoding the desired set of potential protein sequences. The synthesis of degenerate oligonucleotides is well known in the art (see for example, Narang, SA (1983) *Tetrahedron* 39:3; Itakura et al., (1981) *Recombinant DNA, Proc.* 3rd Cleveland Sympos. Macromolecules, ed. AG Walton, Amsterdam: Elsevier pp. 273-289; Itakura et al., (1984) *Annu. Rev. Biochem.* 53:323; Itakura et al., (1984) *Science* 198:1056; Ike et al., (1983) *Nucleic Acid Res.* 11:477). Such techniques have been employed in the directed evolution of other proteins (see, for example, Scott et al., (1990) *Science* 249:386-390; Roberts et al., (1992) *PNAS USA* 89:2429-2433; Devlin et al., (1990) *Science* 249: 404-406; Cwirla et al., (1990) *PNAS USA* 87: 6378-6382; as well as U.S. Patent Nos: 5,223,409, 5,198,346, and 5,096,815).

Alternatively, other forms of mutagenesis may be utilized to generate a combinatorial library. For example, protein homologs (both agonist and antagonist forms) may be generated and isolated from a library by screening using, for example, alanine scanning mutagenesis and the like (Ruf et al., (1994) *Biochemistry* 33:1565-1572; Wang et al., (1994) *J. Biol. Chem.* 269:3095-3099; Balint et al., (1993) *Gene* 137:109-118; Grodberg et al., (1993) *Eur. J. Biochem.* 218:597-601; Nagashima et al., (1993) *J. Biol. Chem.* 268:2888-2892; Lowman et al., (1991) *Biochemistry* 30:10832-10838; and Cunningham et al., (1989) *Science* 244:1081-1085), by linker scanning mutagenesis (Gustin et al., (1993) *Virology* 193:653-660; Brown et al., (1992) *Mol. Cell Biol.* 12:2644-2652; McKnight et al.,

(1982) *Science* 232:316); by saturation mutagenesis (Meyers et al., (1986) *Science* 232:613); by PCR mutagenesis (Leung et al., (1989) *Method Cell Mol Biol* 1:11-19); or by random mutagenesis (Miller et al., (1992) A Short Course in Bacterial Genetics, CSHL Press, Cold Spring Harbor, NY; and Greener et al., (1994) *Strategies in Mol Biol* 7:32-34). Linker scanning mutagenesis, particularly in a combinatorial setting, is an attractive method for identifying truncated forms of proteins that are bioactive.

A wide range of techniques are known in the art for screening gene products of combinatorial libraries made by point mutations and truncations, and for screening cDNA libraries for gene products having a certain property. Such techniques will be generally adaptable for rapid screening of the gene libraries generated by the combinatorial mutagenesis of protein homologs. The most widely used techniques for screening large gene libraries typically comprises cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates relatively easy isolation of the vector encoding the gene whose product was detected. Each of the illustrative assays described below are amenable to high throughput analysis as necessary to screen large numbers of degenerate sequences created by combinatorial mutagenesis techniques.

In an illustrative embodiment of a screening assay, candidate combinatorial gene products are displayed on the surface of a cell and the ability of particular cells or viral particles to bind to the combinatorial gene product is detected in a "panning assay". For instance, the gene library may be cloned into the gene for a surface membrane protein of a bacterial cell (Ladner et al., WO 88/06630; Fuchs et al., (1991) *Bio/Technology* 9:1370-1371; and Goward et al., (1992) *TIBS* 18:136-140), and the resulting fusion protein detected by panning, e.g. using a fluorescently labeled molecule which binds the cell surface protein, e.g. FITC-substrate, to score for potentially functional homologs. Cells may be visually inspected and separated under a fluorescence microscope, or, when the morphology of the cell permits, separated by a fluorescence-activated cell sorter. This method may be used to identify substrates or other polypeptides that can interact with a polypeptide of the invention.

In similar fashion, the gene library may be expressed as a fusion protein on the surface of a viral particle. For instance, in the filamentous phage system, foreign peptide sequences may be expressed on the surface of infectious phage, thereby conferring two

benefits. First, because these phage may be applied to affinity matrices at very high concentrations, a large number of phage may be screened at one time. Second, because each infectious phage displays the combinatorial gene product on its surface, if a particular phage is recovered from an affinity matrix in low yield, the phage may be amplified by another round of infection. The group of almost identical *E. coli* filamentous phages M13, fd, and f1 are most often used in phage display libraries, as either of the phage gIII or gVIII coat proteins may be used to generate fusion proteins without disrupting the ultimate packaging of the viral particle (Ladner et al., PCT publication WO 90/02909; Garrard et al., PCT publication WO 92/09690; Marks et al., (1992) *J. Biol. Chem.* 267:16007-16010; Griffiths et al., (1993) *EMBO J.* 12:725-734; Clackson et al., (1991) *Nature* 352:624-628; and Barbas et al., (1992) *PNAS USA* 89:4457-4461). Other phage coat proteins may be used as appropriate.

The invention also provides for reduction of the polypeptides of the invention to generate mimetics, e.g. peptide or non-peptide agents, which are able to mimic binding of the authentic protein to another cellular partner. Such mutagenic techniques as described above, as well as the thioredoxin system, are also particularly useful for mapping the determinants of a protein which participates in a protein-protein interaction with another protein. To illustrate, the critical residues of a protein which are involved in molecular recognition of a substrate protein may be determined and used to generate peptidomimetics that may bind to the substrate protein. The peptidomimetic may then be used as an inhibitor of the wild-type protein by binding to the substrate and covering up the critical residues needed for interaction with the wild-type protein, thereby preventing interaction of the protein and the substrate. By employing, for example, scanning mutagenesis to map the amino acid residues of a protein which are involved in binding a substrate polypeptide, peptidomimetic compounds may be generated which mimic those residues in binding to the substrate. For instance, non-hydrolyzable peptide analogs of such residues may be generated using benzodiazepine (e.g., see Freidinger et al., in *Peptides: Chemistry and Biology*, G.R. Marshall ed., ESCOM Publisher: Leiden, Netherlands, 1988), azepine (e.g., see Huffman et al., in *Peptides: Chemistry and Biology*, G.R. Marshall ed., ESCOM Publisher: Leiden, Netherlands, 1988), substituted gamma lactam rings (Garvey et al., in *Peptides: Chemistry and Biology*, G.R. Marshall ed., ESCOM Publisher: Leiden, Netherlands, 1988), keto-methylene pseudopeptides (Ewenson et al., (1986) *J. Med. Chem.* 29:295; and Ewenson et al., in *Peptides: Structure and Function* (Proceedings of the 9th

American Peptide Symposium) Pierce Chemical Co. Rockland, IL, 1985), β-turn dipeptide cores (Nagai et al., (1985) *Tetrahedron Lett* 26:647; and Sato et al., (1986) *J Chem Soc Perkin Trans* 1:1231), and β-aminoalcohols (Gordon et al., (1985) *Biochem Biophys Res Commun* 126:419; and Dann et al., (1986) *Biochem Biophys Res Commun* 134:71).

The activity of a polypeptide of the invention may be identified and/or assayed using a variety of methods well known to the skilled artisan. For example, information about the activity of non-essential genes may be assayed by creating a null mutant strain of bacteria expressing a mutant form of, or lacking expression of, a protein of interest. The resulting phenotype of the null mutant strain may provide information about the activity of the mutated gene product. Essential genes may be studied by creating a bacterial strain with a conditional mutation in the gene of interest. The bacterial strain may be grown under permissive and non-permissive conditions and the change in phenotype under the non-permissive conditions may be used to identify and/or assay the activity of the gene product.

In an alternative embodiment, the activity of a protein may be assayed using an appropriate substrate or binding partner or other reagent suitable to test for the suspected activity. For catalytic activity, the assay is typically designed so that the enzymatic reaction produces a detectable signal. For example, mixture of a kinase with a substrate in the presence of ^{32}P will result in incorporation of the ^{32}P into the substrate. The labeled substrate may then be separated from the free ^{32}P and the presence and/or amount of radiolabeled substrate may be detected using a scintillation counter or a phosphorimager. Similar assays may be designed to identify and/or assay the activity of a wide variety of enzymatic activities. Based on the teachings herein, the skilled artisan would readily be able to develop an appropriate assay for a polypeptide of the invention.

In another embodiment, the activity of a polypeptide of the invention may be determined by assaying for the level of expression of RNA and/or protein molecules. Transcription levels may be determined, for example, using Northern blots, hybridization to an oligonucleotide array or by assaying for the level of a resulting protein product. Translation levels may be determined, for example, using Western blotting or by identifying a detectable signal produced by a protein product (e.g., fluorescence, luminescence, enzymatic activity, etc.). Depending on the particular situation, it may be desirable to detect the level of transcription and/or translation of a single gene or of multiple genes.

Alternatively, it may be desirable to measure the overall rate of DNA replication, transcription and/or translation in a cell. In general this may be accomplished by growing the cell in the presence of a detectable metabolite which is incorporated into the resultant DNA, RNA, or protein product. For example, the rate of DNA synthesis may be determined by growing cells in the presence of BrdU which is incorporated into the newly synthesized DNA. The amount of BrdU may then be determined histochemically using an anti-BrdU antibody.

In general, the biological activity of a polypeptide encoded by SEQ ID NO. 2, and possibly other polypeptides of the invention, is peptidyl-tRNA hydrolase, having the gene designation of *pth*. The polypeptide encoded by SEQ ID NO. 2, and possibly other polypeptides of the invention, may be further characterized as being part of the COG category "translation, ribosomal structure and biogenesis", with COG ID No. COG0193. The foregoing annotations were determined in accordance with the procedure described in EXAMPLE 17. This functionality asignment has been confirmed by completion of the X-ray structure of a polypeptide of the invention, as described in more detail below. In one aspect, the present invention contemplates a polypeptide having biological activity, or is a component of a protein complex having biological activity, substantially similar to or identical to peptidyl-tRNA hydrolase. Alternatively, the polypeptide catalyzes, or is a component of a protein complex that catalyzes, a reaction that is substantially the same type of, or is the same as, the reaction catalyzed by peptidyl-tRNA hydrolase. Other biological activities of polypeptides of the invention are described herein, or will be reasonably apparent to those skilled in the art in light of the present disclosure.

During the elongation of a protein, peptidyl-tRNAs can dissociate prematurely from the ribosome, resulting in abortive polypeptide chain termination whereby peptidyl-tRNA dissociates from the ribosome before the hydrolysis of the ester bond ("drop-off"). By cleaving the ester bond linking the tRNA and nascent peptide, peptidyl-tRNA hydrolase (PTH) works to ensure the recycling of all peptidyl-tRNAs produced through the abortion of the translation allowing the tRNA to be charged by the cognate aminoacyl-tRNA synthetase and then reutilized in protein synthesis. Initiator formyl-methionyl-$tRNA_f^{MET}$ is the only aminoacyl-tRNA that is known to be safe from attack by PTH, thus enabling the cell to keep this tRNA intact for translation initiation.

PTH appears to be ubiquitous and has been shown to be essential for growth in *E. coli*, suggesting that its homolog in *P. aeruginosa* is also a potential molecular

therapeutic target for the development of antibacterial agents and other therapeutics. This enzymatic activity is essential for cell viability. In the absence of PTH activity, accumulation of peptidyl-tRNA species in the cell is rapid and protein synthesis becomes inhibited due to the lack of essential tRNA isoacceptors, leading to an inhibition of protein synthesis and eventually to cell death. Expression of very short open reading frames called mini-genes also lead to the inhibition of protein synthesis and an arrest in cell growth because peptidyl-tRNA hydrolase fails to recycle sufficiently rapidly peptidyl-tRNA released from ribosomes at the stop signal in competition with normal termination, causing starvation for essential species of tRNA PTHs have also been recently suggested as facilitator of the splicing in vivo of the largest of the known catalytic RNAs (group II intron) in chloroplasts. Also, erythromycin, carbomycin, and lincosamide antibiotics are believed to function in part by stimulating the dissociation of peptidyl-tRNA from ribosomes, resulting in cell death.

The location of the PTH active site has been proposed to be similar to the one of the zinc-dependant aminopeptidase. The structural superimposition between *E. coli* PTH and the aminopeptidase together with *E. coli* PTH active site mapping has been determinant in a more precise identification of residues important for PTH activity. The three residues N10, H20 and D93 appear to be involved in the enzyme action. The catalytic efficiency of the H20A, N10A and D93A mutants were decreased by a factor of >100 while the K_m value of the substrate (di-acetyl-Lys-$tRNA^{Lys}$) remained unchanged in the case of the N10A mutation, and was increased by a factor of 1.6 in the case of D93A. Those residues make part of a channel at the surface of the enzyme where most of the strictly conserved amino-acids are located.

Site-directed mutagenesis has also been used to identify residues involved in binding of the tRNA moiety of the substrate. The position at the surface of PTH of the 3'-end of peptidyl-tRNAs corresponds to a four asparagines cluster (N10, N21, N68 and N114) located in the catalytic crevice. The receptor site of the 5'-phosphate of peptidyl-tRNAs possibly involves two cationic residues (R133 and K105) bordering each side of the channel. In the *E. coli* PTH, this channel is partially occupied by the C-end of a neighboring molecule. The binding of the three last residues of one PTH molecule to another PTH molecule is proposed to correspond to a product complex of the hydrolase that resembles to a tri-peptide esterified to its tRNA.

An alignment of amino acid sequences of *pth* from 6 pathogens is shown in FIGURE 10, and discussed in greater detail below.

For all of the foregoing reasons, the polypeptides of the present invention are potentially valuable targets for therapeutics and diagnostics.

3. Nucleic Acids of the Invention

One aspect of the invention pertains to isolated nucleic acids of the invention. For example, the present invention contemplates an isolated nucleic acid comprising (a) the nucleotide sequence of SEQ ID NO: 1 or SEQ ID NO: 3, (b) a nucleotide sequence at least 80% identical to SEQ ID NO: 1 or SEQ ID NO: 3, (c) a nucleotide sequence that hybridizes under stringent conditions to SEQ ID NO: 1 or SEQ ID NO: 3, or (d) the complement of the nucleotide sequence of (a), (b) or (c). In certain embodiments, nucleic acids of the invention may be labeled, with for example, a radioactive, chemiluminescent or fluorescent label.

It may be that case that the nucleic acid sequence of SEQ ID NO: 3 differs from that of SEQ ID NO: 1 by one or more nucleic acid residues. SEQ ID NO: 3 is determined experimetally, and SEQ ID NO: 1 obtained as described in EXAMPLE 1. In such a case, the present invention contemplates the specific nucleic acid sequences of SEQ ID NO: 1 and SEQ ID NO: 3, and variants thereof, as well as any differences in the applicable amino acid sequences encoded thereby.

In another aspect, the present invention contemplates an isolated nucleic acid that specifically hybridizes under stringent conditions to at least ten nucleotides of SEQ ID NO: 1 or SEQ ID NO: 3, or the complement thereof, which nucleic acid can specifically detect or amplify SEQ ID NO: 1 or SEQ ID NO: 3, or the complement thereof. In yet another aspect, the present invention contemplates such an isolated nucleic acid comprising a nucleotide sequence encoding a fragment of SEQ ID NO: 2 or SEQ ID NO: 4 at least 8 residues in length. The present invention further contemplates a method of hybridizing an oligonucleotide with a nucleic acid of the invention comprising: (a) providing a single-stranded oligonucleotide at least eight nucleotides in length, the oligonucleotide being complementary to a portion of a nucleic acid of the invention; and (b) contacting the oligonucleotide with a sample comprising a nucleic acid of the acid under conditions that permit hybridization of the oligonucleotide with the nucleic acid of the invention.

Isolated nucleic acids which differ from the nucleic acids of the invention due to degeneracy in the genetic code are also within the scope of the invention. For example, a number of amino acids are designated by more than one triplet. Codons that specify the same amino acid, or synonyms (for example, CAU and CAC are synonyms for histidine) may result in "silent" mutations which do not affect the amino acid sequence of the protein. However, it is expected that DNA sequence polymorphisms that do lead to changes in the amino acid sequences of the polypeptides of the invention will exist. One skilled in the art will appreciate that these variations in one or more nucleotides (from less than 1% up to about 3 or 5% or possibly more of the nucleotides) of the nucleic acids encoding a particular protein of the invention may exist among a given species due to natural allelic variation. Any and all such nucleotide variations and resulting amino acid polymorphisms are within the scope of this invention.

Bias in codon choice within genes in a single species appears related to the level of expression of the protein encoded by that gene. Accordingly, the invention encompasses nucleic acid sequences which have been optimized for improved expression in a host cell by altering the frequency of codon usage in the nucleic acid sequence to approach the frequency of preferred codon usage of the host cell. Due to codon degeneracy, it is possible to optimize the nucleotide sequence without affecting the amino acid sequence of an encoded polypeptide. Accordingly, the instant invention relates to any nucleotide sequence that encodes all or a substantial portion of the amino acid sequence set forth in SEQ ID NO: 2, SEQ ID NO: 4 or other polypeptides of the invention.

The present invention pertains to nucleic acids encoding proteins derived from *P. aeruginosa* and which have amino acid sequences evolutionarily related to a polypeptide of the invention, wherein "evolutionarily related to", refers to proteins having different amino acid sequences which have arisen naturally (e.g. by allelic variance or by differential splicing), as well as mutational variants of the proteins of the invention which are derived, for example, by combinatorial mutagenesis.

Fragments of the polynucleotides of the invention encoding a biologically active portion of the subject polypeptides are also within the scope of the invention. As used herein, a fragment of a nucleic acid of the invention encoding an active portion of a polypeptide of the invention refers to a nucleotide sequence having fewer nucleotides than the nucleotide sequence encoding the full length amino acid sequence of a polypeptide of the invention, for example, SEQ ID NO: 2 or SEQ ID NO: 4, and which encodes a

polypeptide which retains at least a portion of a biological activity of the full-length protein as defined herein, or alternatively, which is functional as a modulator of the biological activity of the full-length protein. For example, such fragments include a polypeptide containing a domain of the full-length protein from which the polypeptide is derived that mediates the interaction of the protein with another molecule (e.g., polypeptide, DNA, RNA, etc.). In another embodiment, the present invention contemplates an isolated nucleic acid that encodes a polypeptide having a biological activity of a protein having the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4, or alternatively biological activity of peptidyl-tRNA hydrolase.

Nucleic acids within the scope of the invention may also contain linker sequences, modified restriction endonuclease sites and other sequences useful for molecular cloning, expression or purification of such recombinant polypeptides.

A nucleic acid encoding a polypeptide of the invention may be obtained from mRNA or genomic DNA from any organism in accordance with protocols described herein, as well as those generally known to those skilled in the art. A cDNA encoding a polypeptide of the invention, for example, may be obtained by isolating total mRNA from an organism, e.g. a bacteria, virus, mammal, etc. Double stranded cDNAs may then be prepared from the total mRNA, and subsequently inserted into a suitable plasmid or bacteriophage vector using any one of a number of known techniques. A gene encoding a polypeptide of the invention may also be cloned using established polymerase chain reaction techniques in accordance with the nucleotide sequence information provided by the invention. In one aspect, the present invention contemplates a method for amplification of a nucleic acid of the invention, or a fragment thereof, comprising: (a) providing a pair of single stranded oligonucleotides, each of which is at least eight nucleotides in length, complementary to sequences of a nucleic acid of the invention, and wherein the sequences to which the oligonucleotides are complementary are at least ten nucleotides apart; and (b) contacting the oligonucleotides with a sample comprising a nucleic acid comprising the nucleic acid of the invention under conditions which permit amplification of the region located between the pair of oligonucleotides, thereby amplifying the nucleic acid.

Another aspect of the invention relates to the use of nucleic acids of the invention in "antisense therapy". As used herein, antisense therapy refers to administration or *in situ* generation of oligonucleotide probes or their derivatives which specifically hybridize or otherwise bind under cellular conditions with the cellular mRNA and/or genomic DNA

encoding one of the polypeptides of the invention so as to inhibit expression of that polypeptide, e.g. by inhibiting transcription and/or translation. The binding may be by conventional base pair complementarity, or, for example, in the case of binding to DNA duplexes, through specific interactions in the major groove of the double helix. In general, antisense therapy refers to the range of techniques generally employed in the art, and includes any therapy which relies on specific binding to oligonucleotide sequences.

An antisense construct of the present invention may be delivered, for example, as an expression plasmid which, when transcribed in the cell, produces RNA which is complementary to at least a unique portion of the mRNA which encodes a polypeptide of the invention. Alternatively, the antisense construct may be an oligonucleotide probe which is generated *ex vivo* and which, when introduced into the cell causes inhibition of expression by hybridizing with the mRNA and/or genomic sequences encoding a polypeptide of the invention. Such oligonucleotide probes may be modified oligonucleotides which are resistant to endogenous nucleases, e.g. exonucleases and/or endonucleases, and are therefore stable *in vivo*. Exemplary nucleic acid molecules for use as antisense oligonucleotides are phosphoramidate, phosphothioate and methylphosphonate analogs of DNA (see also U.S. Patents 5,176,996; 5,264,564; and 5,256,775). Additionally, general approaches to constructing oligomers useful in antisense therapy have been reviewed, for example, by van der Krol et al., (1988) *Biotechniques* 6:958-976; and Stein et al., (1988) *Cancer Res* 48:2659-2668.

In a further aspect, the invention provides double stranded small interfering RNAs (siRNAs), and methods for administering the same. siRNAs decrease or block gene expression. While not wishing to be bound by theory, it is generally thought that siRNAs inhibit gene expression by mediating sequence specific mRNA degradation. RNA interference (RNAi) is the process of sequence-specific, post-transcriptional gene silencing, particularly in animals and plants, initiated by double-stranded RNA (dsRNA) that is homologous in sequence to the silenced gene (Elbashir et al. Nature 2001; 411(6836): 494-8). Accordingly, it is understood that siRNAs and long dsRNAs having substantial sequence identity to all or a portion of SEQ ID NO: 1 or SEQ ID NO: 3 may be used to inhibit the expression of a nucleic acid of the invention, and particularly when the polynucleotide is expressed in a mammalian or plant cell.

The nucleic acids of the invention may be used as diagnostic reagents to detect the presence or absence of the target DNA or RNA sequences to which they specifically bind,

such as for determining the level of expression of a nucleic acid of the invention. In one aspect, the present invention contemplates a method for detecting the presence of a nucleic acid of the invention or a portion thereof in a sample, the method comprising: (a) providing an oligonucleotide at least eight nucleotides in length, the oligonucleotide being complementary to a portion of a nucleic acid of the invention; (b) contacting the oligonucleotide with a sample comprising at least one nucleic acid under conditions that permit hybridization of the oligonucleotide with a nucleic acid comprising a nucleotide sequence complementary thereto; and (c) detecting hybridization of the oligonucleotide to a nucleic acid in the sample, thereby detecting the presence of a nucleic acid of the invention or a portion thereof in the sample. In another aspect, the present invention contemplates a method for detecting the presence of a nucleic acid of the invention or a portion thereof in a sample, the method comprising: (a) providing a pair of single stranded oligonucleotides, each of which is at least eight nucleotides in length, complementary to sequences of a nucleic acid of the invention, and wherein the sequences to which the oligonucleotides are complementary are at least ten nucleotides apart; and (b) contacting the oligonucleotides with a sample comprising at least one nucleic acid under hybridization conditions; (c) amplifying the nucleotide sequence between the two oligonucleotide primers; and (d) detecting the presence of the amplified sequence, thereby detecting the presence of a nucleic acid comprising the nucleic acid of the invention or a portion thereof in the sample.

In another aspect of the invention, the subject nucleic acid is provided in an expression vector comprising a nucleotide sequence encoding a polypeptide of the invention and operably linked to at least one regulatory sequence. It should be understood that the design of the expression vector may depend on such factors as the choice of the host cell to be transformed and/or the type of protein desired to be expressed. The vector's copy number, the ability to control that copy number and the expression of any other protein encoded by the vector, such as antibiotic markers, should be considered.

The subject nucleic acids may be used to cause expression and over-expression of a polypeptide of the invention in cells propagated in culture, e.g. to produce proteins or polypeptides, including fusion proteins or polypeptides.

This invention pertains to a host cell transfected with a recombinant gene in order to express a polypeptide of the invention. The host cell may be any prokaryotic or eukaryotic cell. For example, a polypeptide of the invention may be expressed in bacterial cells, such as *E. coli*, insect cells (baculovirus), yeast, or mammalian cells. In those instances when the

host cell is human, it may or may not be in a live subject. Other suitable host cells are known to those skilled in the art. Additionally, the host cell may be supplemented with tRNA molecules not typically found in the host so as to optimize expression of the polypeptide. Other methods suitable for maximizing expression of the polypeptide will be known to those in the art.

The present invention further pertains to methods of producing the polypeptides of the invention. For example, a host cell transfected with an expression vector encoding a polypeptide of the invention may be cultured under appropriate conditions to allow expression of the polypeptide to occur. The polypeptide may be secreted and isolated from a mixture of cells and medium containing the polypeptide. Alternatively, the polypeptide may be retained cytoplasmically and the cells harvested, lysed and the protein isolated.

A cell culture includes host cells, media and other byproducts. Suitable media for cell culture are well known in the art. The polypeptide may be isolated from cell culture medium, host cells, or both using techniques known in the art for purifying proteins, including ion-exchange chromatography, gel filtration chromatography, ultrafiltration, electrophoresis, and immunoaffinity purification with antibodies specific for particular epitopes of a polypeptide of the invention.

Thus, a nucleotide sequence encoding all or a selected portion of polypeptide of the invention, may be used to produce a recombinant form of the protein via microbial or eukaryotic cellular processes. Ligating the sequence into a polynucleotide construct, such as an expression vector, and transforming or transfecting into hosts, either eukaryotic (yeast, avian, insect or mammalian) or prokaryotic (bacterial cells), are standard procedures. Similar procedures, or modifications thereof, may be employed to prepare recombinant polypeptides of the invention by microbial means or tissue-culture technology.

Expression vehicles for production of a recombinant protein include plasmids and other vectors. For instance, suitable vectors for the expression of a polypeptide of the invention include plasmids of the types: pBR322-derived plasmids, pEMBL-derived plasmids, pEX-derived plasmids, pBTac-derived plasmids and pUC-derived plasmids for expression in prokaryotic cells, such as *E. coli*.

A number of vectors exist for the expression of recombinant proteins in yeast. For instance, YEP24, YIP5, YEP51, YEP52, pYES2, and YRP17 are cloning and expression vehicles useful in the introduction of genetic constructs into *S. cerevisiae* (see, for example, Broach et al., (1983) in *Experimental Manipulation of Gene Expression*, ed. M. Inouye

Academic Press, p. 83). These vectors may replicate in *E. coli* due the presence of the pBR322 ori, and in *S. cerevisiae* due to the replication determinant of the yeast 2 micron plasmid. In addition, drug resistance markers such as ampicillin may be used.

In certain embodiments, mammalian expression vectors contain both prokaryotic sequences to facilitate the propagation of the vector in bacteria, and one or more eukaryotic transcription units that are expressed in eukaryotic cells. The pcDNAI/amp, pcDNAI/neo, pRc/CMV, pSV2gpt, pSV2neo, pSV2-dhfr, pTk2, pRSVneo, pMSG, pSVT7, pko-neo and pHyg derived vectors are examples of mammalian expression vectors suitable for transfection of eukaryotic cells. Some of these vectors are modified with sequences from bacterial plasmids, such as pBR322, to facilitate replication and drug resistance selection in both prokaryotic and eukaryotic cells. Alternatively, derivatives of viruses such as the bovine papilloma virus (BPV-1), or Epstein-Barr virus (pHEBo, pREP-derived and p205) can be used for transient expression of proteins in eukaryotic cells. The various methods employed in the preparation of the plasmids and transformation of host organisms are well known in the art. For other suitable expression systems for both prokaryotic and eukaryotic cells, as well as general recombinant procedures, see *Molecular Cloning A Laboratory Manual*, 2nd Ed., ed. by Sambrook, Fritsch and Maniatis (Cold Spring Harbor Laboratory Press, 1989) Chapters 16 and 17. In some instances, it may be desirable to express the recombinant protein by the use of a baculovirus expression system. Examples of such baculovirus expression systems include pVL-derived vectors (such as pVL1392, pVL1393 and pVL941), pAcUW-derived vectors (such as pAcUW1), and pBlueBac-derived vectors (such as the ß-gal containing pBlueBac III).

In another variation, protein production may be achieved using *in vitro* translation systems. *In vitro* translation systems are, generally, a translation system which is a cell-free extract containing at least the minimum elements necessary for translation of an RNA molecule into a protein. An *in vitro* translation system typically comprises at least ribosomes, tRNAs, initiator methionyl-tRNAMet, proteins or complexes involved in translation, e.g., eIF2, eIF3, the cap-binding (CB) complex, comprising the cap-binding protein (CBP) and eukaryotic initiation factor 4F (eIF4F). A variety of *in vitro* translation systems are well known in the art and include commercially available kits. Examples of *in vitro* translation systems include eukaryotic lysates, such as rabbit reticulocyte lysates, rabbit oocyte lysates, human cell lysates, insect cell lysates and wheat germ extracts. Lysates are commercially available from manufacturers such as Promega Corp., Madison,

Wis.; Stratagene, La Jolla, Calif.; Amersham, Arlington Heights, Ill.; and GIBCO/BRL, Grand Island, N.Y. *In vitro* translation systems typically comprise macromolecules, such as enzymes, translation, initiation and elongation factors, chemical reagents, and ribosomes. In addition, an *in vitro* transcription system may be used. Such systems typically comprise at least an RNA polymerase holoenzyme, ribonucleotides and any necessary transcription initiation, elongation and termination factors. *In vitro* transcription and translation may be coupled in a one-pot reaction to produce proteins from one or more isolated DNAs.

When expression of a carboxy terminal fragment of a polypeptide is desired, i.e. a truncation mutant, it may be necessary to add a start codon (ATG) to the oligonucleotide fragment containing the desired sequence to be expressed. It is well known in the art that a methionine at the N-terminal position may be enzymatically cleaved by the use of the enzyme methionine aminopeptidase (MAP). MAP has been cloned from *E. coli* (Ben-Bassat et al., (1987) *J. Bacteriol.* 169:751-757) and *Salmonella typhimurium* and its *in vitro* activity has been demonstrated on recombinant proteins (Miller et al., (1987) *PNAS USA* *84*:2718-1722). Therefore, removal of an N-terminal methionine, if desired, may be achieved either *in vivo* by expressing such recombinant polypeptides in a host which produces MAP (e.g., *E. coli* or CM89 or *S. cerevisiae*), or *in vitro* by use of purified MAP (e.g., procedure of Miller et al.).

Coding sequences for a polypeptide of interest may be incorporated as a part of a fusion gene including a nucleotide sequence encoding a different polypeptide. The present invention contemplates an isolated nucleic acid comprising a nucleic acid of the invention and at least one heterologous sequence encoding a heterologous peptide linked in frame to the nucleotide sequence of the nucleic acid of the invention so as to encode a fusion protein comprising the heterologous polypeptide. The heterologous polypeptide may be fused to (a) the C-terminus of the polypeptide encoded by the nucleic acid of the invention, (b) the N-terminus of the polypeptide, or (c) the C-terminus and the N-terminus of the polypeptide. In certain instances, the heterologous sequence encodes a polypeptide permitting the detection, isolation, solubilization and/or stabilization of the polypeptide to which it is fused. In still other embodiments, the heterologous sequence encodes a polypeptide selected from the group consisting of a polyHis tag, myc, HA, GST, protein A, protein G, calmodulin-binding peptide, thioredoxin, maltose-binding protein, poly arginine, poly His-Asp, FLAG, a portion of an immunoglobulin protein, and a transcytosis peptide.

Fusion expression systems can be useful when it is desirable to produce an immunogenic fragment of a polypeptide of the invention. For example, the VP6 capsid protein of rotavirus may be used as an immunologic carrier protein for portions of polypeptide, either in the monomeric form or in the form of a viral particle. The nucleic acid sequences corresponding to the portion of a polypeptide of the invention to which antibodies are to be raised may be incorporated into a fusion gene construct which includes coding sequences for a late vaccinia virus structural protein to produce a set of recombinant viruses expressing fusion proteins comprising a portion of the protein as part of the virion. The Hepatitis B surface antigen may also be utilized in this role as well. Similarly, chimeric constructs coding for fusion proteins containing a portion of a polypeptide of the invention and the poliovirus capsid protein may be created to enhance immunogenicity (see, for example, EP Publication NO: 0259149; and Evans et al., (1989) *Nature* 339:385; Huang et al., (1988) *J. Virol.* 62:3855; and Schlienger et al., (1992) *J. Virol.* 66:2).

Fusion proteins may facilitate the expression and/or purification of proteins. For example, a polypeptide of the invention may be generated as a glutathione-S-transferase (GST) fusion protein. Such GST fusion proteins may be used to simplify purification of a polypeptide of the invention, such as through the use of glutathione-derivatized matrices (see, for example, *Current Protocols in Molecular Biology*, eds. Ausubel et al., (N.Y.: John Wiley & Sons, 1991)). In another embodiment, a fusion gene coding for a purification leader sequence, such as a poly-(His)/enterokinase cleavage site sequence at the N-terminus of the desired portion of the recombinant protein, may allow purification of the expressed fusion protein by affinity chromatography using a Ni^{2+} metal resin. The purification leader sequence may then be subsequently removed by treatment with enterokinase to provide the purified protein (e.g., see Hochuli et al., (1987) *J. Chromatography* 411: 177; and Janknecht et al., *PNAS USA* 88:8972).

Techniques for making fusion genes are well known. Essentially, the joining of various DNA fragments coding for different polypeptide sequences is performed in accordance with conventional techniques, employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene may be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments may be carried out using anchor primers which give rise to

complementary overhangs between two consecutive gene fragments which may subsequently be annealed to generate a chimeric gene sequence (see, for example, *Current Protocols in Molecular Biology*, eds. Ausubel et al., John Wiley & Sons: 1992).

The present invention further contemplates a transgenic non-human animal having cells which harbor a transgene comprising a nucleic acid of the invention.

In other embodiments, the invention provides for nucleic acids of the invention immobilized onto a solid surface, including, plates, microtiter plates, slides, beads, particles, spheres, films, strands, precipitates, gels, sheets, tubing, containers, capillaries, pads, slices, etc. The nucleic acids of the invention may be immobilized onto a chip as part of an array. The array may comprise one or more polynucleotides of the invention as described herein. In one embodiment, the chip comprises one or more polynucleotides of the invention as part of an array of *P. aeruginosa* polynucleotide sequences.

In still other embodiments, the invention comprises the sequence of a nucleic acid of the invention in computer readable format. The invention also encompasses a database comprising the sequence of a nucleic acid of the invention.

4. Homology Searching of Nucleotide and Polypeptide Sequences

The nucleotide or amino acid sequences of the invention, including those set forth in the appended Figures, may be used as query sequences against databases such as GenBank, SwissProt, PDB, BLOCKS, and Pima II. These databases contain previously identified and annotated sequences that may be searched for regions of homology (similarity) using BLAST, which stands for Basic Local Alignment Search Tool (Altschul S F (1993) J Mol Evol 36:290-300; Altschul, S F et al (1990) J Mol Biol 215:403-10).

BLAST produces alignments of both nucleotide and amino acid sequences to determine sequence similarity. Because of the local nature of the alignments, BLAST is especially useful in determining exact matches or in identifying homologs which may be of prokaryotic (bacterial) or eukaryotic (animal, fungal or plant) origin. Other algorithms such as the one described in Smith, R. F. and T. F. Smith (1992; Protein Engineering 5:35-51) may be used when dealing with primary sequence patterns and secondary structure gap penalties. In the usual course using BLAST, sequences have lengths of at least 49 nucleotides and no more than 12% uncalled bases (where N is recorded rather than A, C, G, or T).

The BLAST approach, as detailed in Karlin and Altschul (1993; Proc Nat Acad Sci 90:5873-7) searches matches between a query sequence and a database sequence, to evaluate the statistical significance of any matches found, and to report only those matches which satisfy the user-selected threshold of significance. The threshold is typically set at about 10-25 for nucleotides and about 3-15 for peptides.

5. Analysis of Protein Properties

(a) Analysis of Proteins by Mass Spectrometry

Typically, protein characterization by mass spectroscopy first requires protein isolation followed by either chemical or enzymatic digestion of the protein into smaller peptide fragments, whereupon the peptide fragments may be analyzed by mass spectrometry to obtain a peptide map. Mass spectrometry may also be used to identify post-translational modifications (e.g., phosphorylation, etc.) of a polypeptide.

Various mass spectrometers may be used within the present invention. Representative examples include: triple quadrupole mass spectrometers, magnetic sector instruments (magnetic tandem mass spectrometer, JEOL, Peabody, Mass), ionspray mass spectrometers (Bruins et al., Anal Chem. 59:2642-2647, 1987), electrospray mass spectrometers (including tandem, nano- and nano-electrospray tandem) (Fenn et al., Science 246:64-71, 1989), laser desorption time-of-flight mass spectrometers (Karas and Hillenkamp, Anal. Chem. 60:2299-2301, 1988), and a Fourier Transform Ion Cyclotron Resonance Mass Spectrometer (Extrel Corp., Pittsburgh, Mass.).

MALDI ionization is a technique in which samples of interest, in this case peptides and proteins, are co-crystallized with an acidified matrix. The matrix is typically a small molecule that absorbs at a specific wavelength, generally in the ultraviolet (UV) range, and dissipates the absorbed energy thermally. Typically a pulsed laser beam is used to transfer energy rapidly (i.e., a few ns) to the matrix. This transfer of energy causes the matrix to rapidly dissociate from the MALDI plate surface and results in a plume of matrix and the co-crystallized analytes being transferred into the gas phase. MALDI is considered a "soft-ionization" method that typically results in singly-charged species in the gas phase, most often resulting from a protonation reaction with the matrix. MALDI may be coupled in-line with time of flight (TOF) mass spectrometers. TOF detectors are based on the principle that an analyte moves with a velocity proportional to its mass. Analytes of higher mass move slower than analytes of lower mass and thus reach the detector later than lighter

analytes. The present invention contemplates a composition comprising a polypeptide of the invention and a matrix suitable for mass spectrometry. In certain instances, the matrix is a nicotinic acid derivative or a cinnamic acid derivative.

MALDI-TOF MS is easily performed with modern mass spectrometers. Typically the samples of interest, in this case peptides or proteins, are mixed with a matrix and spotted onto a polished stainless steel plate (MALDI plate). Commercially available MALDI plates can presently hold up to 1536 samples per plate. Once spotted with sample, the MALDI sample plate is then introduced into the vacuum chamber of a MALDI mass spectrometer. The pulsed laser is then activated and the mass to charge ratios of the analytes are measured utilizing a time of flight detector. A mass spectrum representing the mass to charge ratios of the peptides/proteins is generated.

As mentioned above, MALDI can be utilized to measure the mass to charge ratios of both proteins and peptides. In the case of proteins, a mixture of intact protein and matrix are co-crystallized on a MALDI target (Karas, M. and Hillenkamp, F. Anal. Chem. 1988, 60 (20) 2299-2301). The spectrum resulting from this analysis is employed to determine the molecular weight of a whole protein. This molecular weight can then be compared to the theoretical weight of the protein and utilized in characterizing the analyte of interest, such as whether or not the protein has undergone post-translational modifications (e.g., example phosphorylation).

In certain embodiments, MALDI mass spectrometry is used for determination of peptide maps of digested proteins. The peptide masses are measured accurately using a MALDI-TOF or a MALDI-Q-Star mass spectrometer, with detection precision down to the low ppm (parts per million) level. The ensemble of the peptide masses observed in a protein digest, such as a tryptic digest, may be used to search protein/DNA databases in a method called peptide mass fingerprinting. In this approach, protein entries in a database are ranked according to the number of experimental peptide masses that match the predicted trypsin digestion pattern. Commercially available software utilizes a search algorithm that provides a scoring scheme based on the size of the databases, the number of matching peptides, and the different peptides. Depending on the number of peptides observed, the accuracy of the measurement, and the size of the genome of the particular species, unambiguous protein identification may be obtained.

Statistical analysis may be performed upon each protein match to determine the validity of the match. Typical constraints include error tolerances within 0.1 Da for

monoisotopic peptide masses, cysteines may be alkylated and searched as carboxyamidomethyl modifications, 0 or 1 missed enzyme cleavages, and no methionine oxidations allowed. Identified proteins may be stored automatically in a relational database with software links to SDS-PAGE images and ligand sequences. Often even a partial peptide map is specific enough for identification of the protein. If no protein match is found, a more error-tolerant search can be used, for example using fewer peptides or allowing a larger margin error with respect to mass accuracy.

Other mass spectroscopy methods such as tandem mass spectrometry or post source decay may be used to obtain sequence information about proteins that cannot be identified by peptide mass mapping, or to confirm the identity of proteins that are tentatively identified by an error-tolerant peptide mass search described above. (Griffin et al, Rapid Commun. Mass. Spectrom. 1995, 9, 1546-51).

(b) Analysis of Proteins by Nuclear Magnetic Resonance (NMR)

NMR may be used to characterize the structure of a polypeptide in accordance with the methods of the invention. In particular, NMR can be used, for example, to determine the three dimensional structure, the conformational state, the aggregation level, the state of protein folding/unfolding or the dynamic properties of a polypeptide. For example, the present invention contemplates a method for determining three dimensional structure information of a polypeptide of the invention, the method comprising: (a) generating a purified isotopically labeled polypeptide of the invention; and (b) subjecting the polypeptide to NMR spectroscopic analysis, thereby determining information about its three dimensional structure.

Interaction between a polypeptide and another molecule can also be monitored using NMR. Thus, the invention encompasses methods for detecting, designing and characterizing interactions between a polypeptide and another molecule, including polypeptides, nucleic acids and small molecules, utilizing NMR techniques. For example, the present invention contemplates a method for determining three dimensional structure information of a polypeptide of the invention, or a fragment thereof, while the polypeptide is complexed with another molecule, the method comprising: (a) generating a purified isotopically labeled polypeptide of the invention, or a fragment thereof; (b) forming a complex between the polypeptide and the other molecule; and (c) subjecting the complex to NMR spectroscopic analysis, thereby determining information about the three dimensional structure of the polypeptide. In another aspect, the present invention contemplates a

method for identifying compounds that bind to a polypeptide of the invention, or a fragment thereof, the method comprising: (a) generating a first NMR spectrum of an isotopically labeled polypeptide of the invention, or a fragment thereof; (b) exposing the polypeptide to one or more chemical compounds; (c) generating a second NMR spectrum of the polypeptide which has been exposed to one or more chemical compounds; and (d) comparing the first and second spectra to determine differences between the first and the second spectra, wherein the differences are indicative of one or more compounds that have bound to the polypeptide.

Briefly, the NMR technique involves placing the material to be examined (usually in a suitable solvent) in a powerful magnetic field and irradiating it with radio frequency (rf) electromagnetic radiation. The nuclei of the various atoms will align themselves with the magnetic field until energized by the rf radiation. They then absorb this resonant energy and re-radiate it at a frequency dependent on i) the type of nucleus and ii) its atomic environment. Moreover, resonant energy may be passed from one nucleus to another, either through bonds or through three-dimensional space, thus giving information about the environment of a particular nucleus and nuclei in its vicinity.

However, it is important to recognize that not all nuclei are NMR active. Indeed, not all isotopes of the same element are active. For example, whereas "ordinary" hydrogen, ^{1}H, is NMR active, heavy hydrogen (deuterium), ^{2}H, is not active in the same way. Thus, any material that normally contains ^{1}H hydrogen may be rendered "invisible" in the hydrogen NMR spectrum by replacing all or almost all the ^{1}H hydrogens with ^{2}H. It is for this reason that NMR spectroscopic analyses of water-soluble materials frequently are performed in $^{2}H_2O$ (or deuterium) to eliminate the water signal.

Conversely, "ordinary" carbon, ^{12}C, is NMR inactive whereas the stable isotope, ^{13}C, present to about 1% of total carbon in nature, is active. Similarly, while "ordinary" nitrogen, ^{14}N, is NMR active, it has undesirable properties for NMR and resonates at a different frequency from the stable isotope ^{15}N, present to about 0.4% of total nitrogen in nature.

By labeling proteins with ^{15}N and $^{15}N/^{13}C$, it is possible to conduct analytical NMR of macromolecules with weights of 15 kD and 40 kD, respectively. More recently, partial deuteration of the protein in addition to ^{13}C- and ^{15}N-labeling has increased the possible weight of proteins and protein complexes for NMR analysis still further, to approximately 60-70 kD. See Shan et al., J. Am. Chem.Soc., 118:6570-6579 (1996); L.E. Kay, Methods

Enzymol., 339:174-203 (2001); and K.H. Gardner & L.E. Kay, Annu Rev Biophys Biomol Struct., 27:357-406 (1998); and references cited therein.

Isotopic substitution may be accomplished by growing a bacterium or yeast or other type of cultured cells, transformed by genetic engineering to produce the protein of choice, in a growth medium containing ^{13}C-, ^{15}N- and/or ^{2}H-labeled substrates. In certain instances, bacterial growth media consists of ^{13}C-labeled glucose and/or ^{15}N-labeled ammonium salts dissolved in D_2O where necessary. Kay, L. et al., Science, 249:411 (1990) and references therein and Bax, A., J. Am. Chem. Soc., 115, 4369 (1993). More recently, isotopically labeled media especially adapted for the labeling of bacterially produced macromolecules have been described. See U.S. Pat. No. 5,324,658.

The goal of these methods has been to achieve universal and/or random isotopic enrichment of all of the amino acids of the protein. By contrast, other methods allow only certain residues to be relatively enriched in ^{1}H, ^{2}H, ^{13}C and ^{15}N. For example, Kay et al., J. Mol. Biol., 263, 627-636 (1996) and Kay et al., J. Am. Chem. Soc., 119, 7599-7600 (1997) have described methods whereby isoleucine, alanine, valine and leucine residues in a protein may be labeled with ^{2}H, ^{13}C and ^{15}N, and may be specifically labeled with ^{1}H at the terminal methyl position. In this way, study of the proton-proton interactions between some amino acids may be facilitated. Similarly, a cell-free system has been described by Yokoyama et al., J. Biomol. NMR, 6(2), 129-134 (1995), wherein a transcription-translation system derived from *E. coli* was used to express human Ha-Ras protein incorporating ^{15}N into serine and/or aspartic acid.

Techniques for producing isotopically labeled proteins and macromolecules, such as glycoproteins, in mammalian or insect cells have been described. See U.S. Pat. Nos. 5,393,669 and 5,627,044; Weller, C. T., Biochem., 35, 8815-23 (1996) and Lustbader, J. W., J.Biomol. NMR, 7, 295-304 (1996). Other methods for producing polypeptides and other molecules with labels appropriate for NMR are known in the art.

The present invention contemplates using a variety of solvents which are appropriate for NMR. For ^{1}H NMR, a deuterium lock solvent may be used. Exemplary deuterium lock solvents include acetone (CD_3COCD_3), chloroform ($CDCl_3$), dichloro methane (CD_2Cl_2), methylnitrile (CD_3CN), benzene (C_6D_6), water (D_2O), diethylether ($(CD_3CD_2)_2O$), dimethylether ($(CD_3)_2O$), N,N-dimethylformamide ($(CD_3)_2NCDO$), dimethyl sulfoxide (CD_3SOCD_3), ethanol (CD_3CD_2OD), methanol (CD_3OD), tetrahydrofuran (C_4D_8O), toluene ($C_6D_5CD_3$), pyridine (C_5D_5N) and cyclohexane (C_6H_{12}).

For example, the present invention contemplates a composition comprising a polypeptide of the invention and a deuterium lock solvent.

The 2-dimensional ^{1}H-^{15}N HSQC (Heteronuclear Single Quantum Correlation) spectrum provides a diagnostic fingerprint of conformational state, aggregation level, state of protein folding, and dynamic properties of a polypeptide (Yee et al, PNAS 99, 1825-30 (2002)). Polypeptides in aqueous solution usually populate an ensemble of 3-dimensional structures which can be determined by NMR. When the polypeptide is a stable globular protein or domain of a protein, then the ensemble of solution structures is one of very closely related conformations. In this case, one peak is expected for each non-proline residue with a dispersion of resonance frequencies with roughly equal intensity. Additional pairs of peaks from side-chain NH_2 groups are also often observed, and correspond to the approximate number of Gln and Asn residues in the protein. This type of HSQC spectra usually indicates that the protein is amenable to structure determination by NMR methods.

If the HSQC spectrum shows well-dispersed peaks but there are either too few or too many in number, and/or the peak intensities differ throughout the spectrum, then the protein likely does not exist in a single globular conformation. Such spectral features are indicative of conformational heterogeneity with slow or nonexistent inter-conversion between states (too many peaks) or the presence of dynamic processes on an intermediate timescale that can broaden and obscure the NMR signals. Proteins with this type of spectrum can sometimes be stabilized into a single conformation by changing either the protein construct, the solution conditions, temperature or by binding of another molecule.

The ^{1}H-^{15}N HSQC can also indicate whether a protein has formed large nonspecific aggregates or has dynamic properties. Alternatively, proteins that are largely unfolded, e.g., having very little regular secondary structure, result in ^{1}H-^{15}N HSQC spectra in which the peaks are all very narrow and intense, but have very little spectral dispersion in the ^{15}N-dimension. This reflects the fact that many or most of the amide groups of amino acids in unfolded polypeptides are solvent exposed and experience similar chemical environments resulting in similar ^{1}H chemical shifts.

The use of the ^{1}H-^{15}N HSQC, can thus allow the rapid characterization of the conformational state, aggregation level, state of protein folding, and dynamic properties of a polypeptide. Additionally, other 2D spectra such as ^{1}H-^{13}C HSQC, or HNCO spectra can also be used in a similar manner. Further use of the ^{1}H-^{15}N HSQC combined with relaxation measurements can reveal the molecular rotational correlation time and dynamic

properties of polypeptides. The rotational correlation time is proportional to size of the protein and therefore can reveal if it forms specific homo-oligomers such as homodimers, homotetramers, etc.

The structure of stable globular proteins can be determined through a series of well-described procedures. For a general review of structure determination of globular proteins in solution by NMR spectroscopy, see Wüthrich, Science 243: 45-50 (1989). See also, Billeter et al., J. Mol. Biol. 155: 321-346 (1982). Current methods for structure determination usually require the complete or nearly complete sequence-specific assignment of ^{1}H-resonance frequencies of the protein and subsequent identification of approximate inter-hydrogen distances (from nuclear Overhauser effect (NOE) spectra) for use in restrained molecular dynamics calculations of the protein conformation. One approach for the analysis of NMR resonance assignments was first outlined by Wüthrich, Wagner and co-workers (Wüthrich, "NMR or proteins and nucleic acids" Wiley, New York, New York (1986); Wüthrich, Science 243: 45-50 (1989); Billeter et al., J. Mol. Biol. 155: 321-346 (1982)). Newer methods for determining the structures of globular proteins include the use of residual dipolar coupling restraints (Tian et al., J Am Chem Soc. 2001 Nov 28;123(47):11791-6; Bax et al, Methods Enzymol. 2001;339:127-74) and empirically derived conformational restraints (Zweckstetter & Bax, J Am Chem Soc. 2001 Sep 26;123(38):9490-1). It has also been shown that it may be possible to determine structures of globular proteins using only un-assigned NOE measurements. NMR may also be used to determine ensembles of many inter-converting, unfolded conformations (Choy and Forman-Kay, J Mol Biol. 2001 May 18;308(5):1011-32).

NMR analysis of a polypeptide in the presence and absence of a test compound (e.g., a polypeptide, nucleic acid or small molecule) may be used to characterize interactions between a polypeptide and another molecule. Because the ^{1}H-^{15}N HSQC spectrum and other simple 2D NMR experiments can be obtained very quickly (on the order of minutes depending on protein concentration and NMR instrumentation), they are very useful for rapidly testing whether a polypeptide is able to bind to another molecule. Changes in the resonance frequency (in one or both dimensions) of one or more peaks in the HSQC spectrum indicate an interaction with another molecule. Often only a subset of the peaks will have changes in resonance frequency upon binding to anther molecule, allowing one to map onto the structure those residues directly involved in the interaction or involved in conformational changes as a result of the interaction. If the interacting

molecule is relatively large (protein or nucleic acid) the peak widths will also broaden due to the increased rotational correlation time of the complex. In some cases the peaks involved in the interaction may actually disappear from the NMR spectrum if the interacting molecule is in intermediate exchange on the NMR timescale (i.e., exchanging on and off the polypeptide at a frequency that is similar to the resonance frequency of the monitored nuclei).

To facilitate the acquisition of NMR data on a large number of compounds (e.g., a library of synthetic or naturally-occurring small organic compounds), a sample changer may be employed. Using the sample changer, a larger number of samples, numbering 60 or more, may be run unattended. To facilitate processing of the NMR data, computer programs are used to transfer and automatically process the multiple one-dimensional NMR data.

In one embodiment, the invention provides a screening method for identifying small molecules capable of interacting with a polypeptide of the invention. In one example, the screening process begins with the generation or acquisition of either a T_2-filtered or a diffusion-filtered one-dimensional proton spectrum of the compound or mixture of compounds. Means for generating T_2-filtered or diffusion-filtered one-dimensional proton spectra are well known in the art (see, e.g., S. Meiboom and D. Gill, Rev. Sci. Instrum. 29:688(1958), S. J. Gibbs and C. S. Johnson, Jr. J. Main. Reson. 93:395-402 (1991) and A. S. Altieri, et al. J. Am. Chem. Soc. 117: 7566-7567 (1995)).

Following acquisition of the first spectrum for the molecules, the ^{15}N- or ^{13}C-labeled polypeptide is exposed to one or more molecules. Where more than one test compound is to be tested simultaneously, it is preferred to use a library of compounds such as a plurality of small molecules. Such molecules are typically dissolved in perdeuterated dimethylsulfoxide. The compounds in the library may be purchased from vendors or created according to desired needs.

Individual compounds may be selected inter alia on the basis of size and molecular diversity for maximizing the possibility of discovering compounds that interact with widely diverse binding sites of a polypeptide of the invention.

The NMR screening process of the present invention utilizes a range of test compound concentrations, e.g., from about 0.05 to about 1.0 mM. At those exemplary concentrations, compounds which are acidic or basic may significantly change the pH of buffered protein solutions. Chemical shifts are sensitive to pH changes as well as direct

binding interactions, and false-positive chemical shift changes, which are not the result of test compound binding but of changes in pH, may therefore be observed. It may therefore be necessary to ensure that the pH of the buffered solution does not change upon addition of the test compound.

Following exposure of the test compounds to a polypeptide (e.g., the target molecule for the experiment) a second one-dimensional T_2- or diffusion-filtered spectrum is generated. For the T_2-filtered approach, that second spectrum is generated in the same manner as set forth above. The first and second spectra are then compared to determine whether there are any differences between the two spectra. Differences in the one-dimensional T_2-filtered spectra indicate that the compound is binding to, or otherwise interacting with, the target molecule. Those differences are determined using standard procedures well known in the art. For the diffusion-filtered method, the second spectrum is generated by looking at the spectral differences between low and high gradient strengths--thus selecting for those compounds whose diffusion rates are comparable to that observed in the absence of target molecule.

To discover additional molecules that bind to the protein, molecules are selected for testing based on the structure/activity relationships from the initial screen and/or structural information on the initial leads when bound to the protein. By way of example, the initial screening may result in the identification of compounds, all of which contain an aromatic ring. The second round of screening would then use other aromatic molecules as the test compounds.

In another embodiment, the methods of the invention utilize a process for detecting the binding of one ligand to a polypeptide in the presence of a second ligand. In accordance with this embodiment, a polypeptide is bound to the second ligand before exposing the polypeptide to the test compounds.

For more information on NMR methods encompassed by the present invention, see also: U.S. Patent Nos. 5,668,734; 6,194,179; 6,162,627; 6,043,024; 5,817,474; 5,891,642; 5,989,827; 5,891,643; 6,077,682; WO 00/05414; WO 99/22019; Cavanagh, et al., Protein NMR Spectroscopy, Principles and Practice, 1996, Academic Press; Clore, et al., NMR of Proteins. In Topics in Molecular and Structural Biology, 1993, S. Neidle, Fuller, W., and Cohen, J.S., eds., Macmillan Press, Ltd., London; and Christendat et al., Nature Structural Biology 7: 903-909 (2000).

(c) Analysis of Proteins by X-ray Crystallography

(i) X-ray Structure Determination

Exemplary methods for obtaining the three dimensional structure of the crystalline form of a molecule or complex are described herein and, in view of this specification, variations on these methods will be apparent to those skilled in the art (see Ducruix and Geige 1992, IRL Press, Oxford, England).

A variety of methods involving x-ray crystallography are contemplated by the present invention. For example, the present invention contemplates producing a crystallized polypeptide of the invention, or a fragment thereof, by: (a) introducing into a host cell an expression vector comprising a nucleic acid encoding for a polypeptide of the invention, or a fragment thereof; (b) culturing the host cell in a cell culture medium to express the polypeptide or fragment; (c) isolating the polypeptide or fragment from the cell culture; and (d) crystallizing the polypeptide or fragment thereof. Alternatively, the present invention contemplates determining the three dimensional structure of a crystallized polypeptide of the invention, or a fragment thereof, by: (a) crystallizing a polypeptide of the invention, or a fragment thereof, such that the crystals will diffract x-rays to a resolution of 3.5 Å or better; and (b) analyzing the polypeptide or fragment by x-ray diffraction to determine the three-dimensional structure of the crystallized polypeptide.

X-ray crystallography techniques generally require that the protein molecules be available in the form of a crystal. Crystals may be grown from a solution containing a purified polypeptide of the invention, or a fragment thereof (e.g., a stable domain), by a variety of conventional processes. These processes include, for example, batch, liquid, bridge, dialysis, vapour diffusion (e.g., hanging drop or sitting drop methods). (See for example, McPherson, 1982 John Wiley, New York; McPherson, 1990, Eur. J. Biochem. 189: 1-23; Webber. 1991, Adv. Protein Chem. 41:1-36).

In certain embodiments, native crystals of the invention may be grown by adding precipitants to the concentrated solution of the polypeptide. The precipitants are added at a concentration just below that necessary to precipitate the protein. Water may be removed by controlled evaporation to produce precipitating conditions, which are maintained until crystal growth ceases.

The formation of crystals is dependent on a number of different parameters, including pH, temperature, protein concentration, the nature of the solvent and precipitant, as well as the presence of added ions or ligands to the protein. In addition, the sequence of the polypeptide being crystallized will have a significant affect on the success of obtaining

crystals. Many routine crystallization experiments may be needed to screen all these parameters for the few combinations that might give crystal suitable for x-ray diffraction analysis (See, for example, Jancarik, J & Kim, S.H., J. Appl. Cryst. 1991 24: 409-411).

Crystallization robots may automate and speed up the work of reproducibly setting up large number of crystallization experiments. Once some suitable set of conditions for growing the crystal are found, variations of the condition may be systematically screened in order to find the set of conditions which allows the growth of sufficiently large, single, well ordered crystals. In certain instances, a polypeptide of the invention is co-crystallized with a compound that stabilizes the polypeptide.

A number of methods are available to produce suitable radiation for x-ray diffraction. For example, x-ray beams may be produced by synchrotron rings where electrons (or positrons) are accelerated through an electromagnetic field while traveling at close to the speed of light. Because the admitted wavelength may also be controlled, synchrotrons may be used as a tunable x-ray source (Hendrickson WA., Trends Biochem Sci 2000 Dec; 25(12):637-43). For less conventional Laue diffraction studies, polychromatic x-rays covering a broad wavelength window are used to observe many diffraction intensities simultaneously (Stoddard, B. L., Curr. Opin. Struct Biol 1998 Oct; 8(5):612-8). Neutrons may also be used for solving protein crystal structures (Gutberlet T, Heinemann U & Steiner M., Acta Crystallogr D 2001;57: 349-54).

Before data collection commences, a protein crystal may be frozen to protect it from radiation damage. A number of different cryo-protectants may be used to assist in freezing the crystal, such as methyl pentanediol (MPD), isopropanol, ethylene glycol, glycerol, formate, citrate, mineral oil, or a low-molecular-weight polyethylene glycol (PEG). The present invention contemplates a composition comprising a polypeptide of the invention and a cryo-protectant. As an alternative to freezing the crystal, the crystal may also be used for diffraction experiments performed at temperatures above the freezing point of the solution. In these instances, the crystal may be protected from drying out by placing it in a narrow capillary of a suitable material (generally glass or quartz) with some of the crystal growth solution included in order to maintain vapour pressure.

X-ray diffraction results may be recorded by a number of ways know to one of skill in the art. Examples of area electronic detectors include charge coupled device detectors, multi-wire area detectors and phosphoimager detectors (Amemiya, Y, 1997. Methods in Enzymology, Vol. 276. Academic Press, San Diego, pp. 233-243; Westbrook, E. M.,

Naday, I. 1997. Methods in Enzymology, Vol. 276. Academic Press, San Diego, pp. 244-268; 1997. Kahn, R. & Fourme, R. Methods in Enzymology, Vol. 276. Academic Press, San Diego, pp. 268-286).

A suitable system for laboratory data collection might include a Bruker AXS Proteum R system, equipped with a copper rotating anode source, Confocal Max-Flux™ optics and a SMART 6000 charge coupled device detector. Collection of x-ray diffraction patterns are well documented by those skilled in the art (See, for example, Ducruix and Geige, 1992, IRL Press, Oxford, England).

The theory behind diffraction by a crystal upon exposure to x-rays is well known. Because phase information is not directly measured in the diffraction experiment, and is needed to reconstruct the electron density map, methods that can recover this missing information are required. One method of solving structures *ab initio* are the real / reciprocal space cycling techniques. Suitable real / reciprocal space cycling search programs include shake-and-bake (Weeks CM, DeTitta GT, Hauptman HA, Thuman P, Miller R Acta Crystallogr A 1994; V50: 210-20).

Other methods for deriving phases may also be needed. These techniques generally rely on the idea that if two or more measurements of the same reflection are made where strong, measurable, differences are attributable to the characteristics of a small subset of the atoms alone, then the contributions of other atoms can be, to a first approximation, ignored, and positions of these atoms may be determined from the difference in scattering by one of the above techniques. Knowing the position and scattering characteristics of those atoms, one may calculate what phase the overall scattering must have had to produce the observed differences.

One version of this technique is isomorphous replacement technique, which requires the introduction of new, well ordered, x-ray scatterers into the crystal. These additions are usually heavy metal atoms, (so that they make a significant difference in the diffraction pattern); and if the additions do not change the structure of the molecule or of the crystal cell, the resulting crystals should be isomorphous. Isomorphous replacement experiments are usually performed by diffusing different heavy-metal metals into the channels of a pre-existing protein crystal. Growing the crystal from protein that has been soaked in the heavy atom is also possible (Petsko, G.A., 1985. Methods in Enzymology, Vol. 114. Academic Press, Orlando, pp. 147-156). Alternatively, the heavy atom may also be reactive and attached covalently to exposed amino acid side chains (such as the sulfur atom of cysteine)

or it may be associated through non-covalent interactions. It is sometimes possible to replace endogenous light metals in metallo-proteins with heavier ones, e.g., zinc by mercury, or calcium by samarium (Petsko, G.A., 1985. Methods in Enzymology, Vol. 114. Academic Press, Orlando, pp. 147-156). Exemplary sources for such heavy compounds include, without limitation, sodium bromide, sodium selenate, trimethyl lead acetate, mercuric chloride, methyl mercury acetate, platinum tetracyanide, platinum tetrachloride, nickel chloride, and europium chloride.

A second technique for generating differences in scattering involves the phenomenon of anomalous scattering. X-rays that cause the displacement of an electron in an inner shell to a higher shell are subsequently rescattered, but there is a time lag that shows up as a phase delay. This phase delay is observed as a (generally quite small) difference in intensity between reflections known as Friedel mates that would be identical if no anomalous scattering were present. A second effect related to this phenomenon is that differences in the intensity of scattering of a given atom will vary in a wavelength dependent manner, given rise to what are known as dispersive differences. In principle anomalous scattering occurs with all atoms, but the effect is strongest in heavy atoms, and may be maximized by using x-rays at a wavelength where the energy is equal to the difference in energy between shells. The technique therefore requires the incorporation of some heavy atom much as is needed for isomorphous replacement, although for anomalous scattering a wider variety of atoms are suitable, including lighter metal atoms (copper, zinc, iron) in metallo-proteins. One method for preparing a protein for anomalous scattering involves replacing the methionine residues in whole or in part with selenium containing seleno-methionine. Soaks with halide salts such as bromides and other non-reactive ions may also be effective (Dauter Z, Li M, Wlodawer A., Acta Crystallogr D 2001; 57: 239-49).

In another process, known as multiple anomalous scattering or MAD, two to four suitable wavelengths of data are collected. (Hendrickson, W.A. and Ogata, C.M. 1997 Methods in Enzymology 276, 494 – 523). Phasing by various combinations of single and multiple isomorphous and anomalous scattering are possible too. For example, SIRAS (single isomorphous replacement with anomalous scattering) utilizes both the isomorphous and anomalous differences for one derivative to derive phases. More traditionally, several different heavy atoms are soaked into different crystals to get sufficient phase information from isomorphous differences while ignoring anomalous scattering, in the technique known

as multiple isomorphous replacement (MIR) (Petsko, G.A., 1985. Methods in Enzymology, Vol. 114. Academic Press, Orlando, pp. 147-156).

Additional restraints on the phases may be derived from density modification techniques. These techniques use either generally known features of electron density distribution or known facts about that particular crystal to improve the phases. For example, because protein regions of the crystal scatter more strongly than solvent regions, solvent flattening/flipping may be used to adjust phases to make solvent density a uniform flat value (Zhang, K. Y. J., Cowtan, K. and Main, P. Methods in Enzymology 277, 1997 Academic Press, Orlando pp 53-64). If more than one molecule of the protein is present in the asymmetric unit, the fact that the different molecules should be virtually identical may be exploited to further reduce phase error using non-crystallographic symmetry averaging (Villieux, F. M. D. and Read, R. J. Methods in Enzymology 277, 1997 Academic Press, Orlando pp18-52). Suitable programs for performing these processes include DM and other programs of the CCP4 suite (Collaborative Computational Project, Number 4. 1994. Acta Cryst. D50, 760-763) and CNX.

The unit cell dimensions, symmetry, vector amplitude and derived phase information can be used in a Fourier transform function to calculate the electron density in the unit cell, i.e., to generate an experimental electron density map. This may be accomplished using programs of the CNX or CCP4 packages. The resolution is measured in Ångstrom (Å) units, and is closely related to how far apart two objects need to be before they can be reliably distinguished. The smaller this number is, the higher the resolution and therefore the greater the amount of detail that can be seen. Preferably, crystals of the invention diffract x-rays to a resolution of better than about 4.0, 3.5, 3.0, 2.5, 2.0, 1.5, 1.0, 0.5 Å or better.

As used herein, the term "modeling" includes the quantitative and qualitative analysis of molecular structure and/or function based on atomic structural information and interaction models. The term "modeling" includes conventional numeric-based molecular dynamic and energy minimization models, interactive computer graphic models, modified molecular mechanics models, distance geometry and other structure-based constraint models.

Model building may be accomplished by either the crystallographer using a computer graphics program such as TURBO or O (Jones, TA. et al., Acta Crystallogr. A47, 100-119, 1991) or, under suitable circumstances, by using a fully automated model building

program, such as wARP (Anastassis Perrakis, Richard Morris & Victor S. Lamzin; Nature Structural Biology, May 1999 Volume 6 Number 5 pp 458 – 463) or MAID (Levitt, D. G., Acta Crystallogr. D 2001 V57: 1013-9). This structure may be used to calculate model-derived diffraction amplitudes and phases. The model-derived and experimental diffraction amplitudes may be compared and the agreement between them can be described by a parameter referred to as R-factor. A high degree of correlation in the amplitudes corresponds to a low R-factor value, with 0.0 representing exact agreement and 0.59 representing a completely random structure. Because the R-factor may be lowered by introducing more free parameters into the model, an unbiased, cross-correlated version of the R-factor known as the R-free gives a more objective measure of model quality. For the calculation of this parameter a subset of reflections (generally around 10%) are set aside at the beginning of the refinement and not used as part of the refinement target. These reflections are then compared to those predicted by the model (Kleywegt GJ, Brunger AT, Structure 1996 Aug 15;4(8):897-904).

The model may be improved using computer programs that maximize the probability that the observed data was produced from the predicted model, while simultaneously optimizing the model geometry. For example, the CNX program may be used for model refinement, as can the XPLOR program (1992, Nature 355:472-475, G.N. Murshudov, A.A.Vagin and E.J.Dodson, (1997) Acta Cryst. D 53, 240-255). In order to maximize the convergence radius of refinement, simulated annealing refinement using torsion angle dynamics may be employed in order to reduce the degrees of freedom of motion of the model (Adams PD, Pannu NS, Read RJ, Brunger AT., Proc Natl Acad Sci U S A 1997 May 13;94(10):5018-23). Where experimental phase information is available (e.g. where MAD data was collected) Hendrickson-Lattman phase probability targets may be employed. Isotropic or anisotropic domain, group or individual temperature factor refinement, may be used to model variance of the atomic position from its mean. Well defined peaks of electron density not attributable to protein atoms are generally modeled as water molecules. Water molecules may be found by manual inspection of electron density maps, or with automatic water picking routines. Additional small molecules, including ions, cofactors, buffer molecules or substrates may be included in the model if sufficiently unambiguous electron density is observed in a map.

In general, the R-free is rarely as low as 0.15 and may be as high as 0.35 or greater for a reasonably well-determined protein structure. The residual difference is a

consequence of approximations in the model (inadequate modeling of residual structure in the solvent, modeling atoms as isotropic Gaussian spheres, assuming all molecules are identical rather than having a set of discrete conformers, etc.) and errors in the data (Lattman EE., Proteins 1996; 25: i-ii). In refined structures at high resolution, there are usually no major errors in the orientation of individual residues, and the estimated errors in atomic positions are usually around 0.1 - 0.2 up to 0.3 Å.

The three dimensional structure of a new crystal may be modeled using molecular replacement. The term "molecular replacement" refers to a method that involves generating a preliminary model of a molecule or complex whose structure coordinates are unknown, by orienting and positioning a molecule whose structure coordinates are known within the unit cell of the unknown crystal, so as best to account for the observed diffraction pattern of the unknown crystal. Phases may then be calculated from this model and combined with the observed amplitudes to give an approximate Fourier synthesis of the structure whose coordinates are unknown. This, in turn, can be subject to any of the several forms of refinement to provide a final, accurate structure of the unknown crystal. Lattman, E., "Use of the Rotation and Translation Functions", in Methods in Enzymology, 115, pp. 55-77 (1985); M. G. Rossmann, ed., "The Molecular Replacement Method", Int. Sci. Rev. Ser., No. 13, Gordon & Breach, New York, (1972).

Commonly used computer software packages for molecular replacement are CNX, X-PLOR (Brunger 1992, Nature 355: 472-475), AMoRE (Navaza, 1994, Acta Crystallogr. A50:157-163), the CCP4 package, the MERLOT package (P.M.D. Fitzgerald, J. Appl. Cryst., Vol. 21, pp. 273-278, 1988) and XTALVIEW (McCree et al (1992) J. Mol. Graphics 10: 44-46). The quality of the model may be analyzed using a program such as PROCHECK or 3D-Profiler (Laskowski et al 1993 J. Appl. Cryst. 26:283-291; Luthy R. et al, Nature 356: 83-85, 1992; and Bowie, J.U. et al, Science 253: 164-170, 1991).

Homology modeling (also known as comparative modeling or knowledge-based modeling) methods may also be used to develop a three dimensional model from a polypeptide sequence based on the structures of known proteins. The method utilizes a computer model of a known protein, a computer representation of the amino acid sequence of the polypeptide with an unknown structure, and standard computer representations of the structures of amino acids. This method is well known to those skilled in the art (Greer, 1985, Science 228, 1055; Bundell et al 1988, Eur. J. Biochem. 172, 513; Knighton et al., 1992, Science 258:130-135, http://biochem.vt.edu/courses/-modeling/homology.htn).

Computer programs that can be used in homology modeling are QUANTA and the Homology module in the Insight II modeling package distributed by Molecular Simulations Inc, or MODELLER (Rockefeller University, www.iucr.ac.uk/sinris-top/logical/prg-modeller.html).

Once a homology model has been generated it is analyzed to determine its correctness. A computer program available to assist in this analysis is the Protein Health module in QUANTA which provides a variety of tests. Other programs that provide structure analysis along with output include PROCHECK and 3D-Profiler (Luthy R. et al, Nature 356: 83-85, 1992; and Bowie, J.U. et al, Science 253: 164-170, 1991). Once any irregularities have been resolved, the entire structure may be further refined.

Other molecular modeling techniques may also be employed in accordance with this invention. See, e.g., Cohen, N. C. *et al*, J. Med. Chem., 33, pp. 883-894 (1990). See also, Navix, M. A. and M. A. Marko, Current Opinions in Structural Biology, 2, pp. 202-210 (1992).

Under suitable circumstances, the entire process of solving a crystal structure may be accomplished in an automated fashion by a system such as ELVES (http://ucxray.berkeley.edu/~jamesh/elves/index.html) with little or no user intervention.

(ii) X-ray Structure

The present invention provides methods for determining some or all of the structural coordinates for amino acids of a polypeptide of the invention, or a complex thereof.

In another aspect, the present invention provides methods for identifying a druggable region of a polypeptide of the invention. For example, one such method includes: (a) obtaining crystals of a polypeptide of the invention or a fragment thereof such that the three dimensional structure of the crystallized protein can be determined to a resolution of 3.5 Å or better; (b) determining the three dimensional structure of the crystallized polypeptide or fragment using x-ray diffraction; and (c) identifying a druggable region of a polypeptide of the invention based on the three-dimensional structure of the polypeptide or fragment.

A three dimensional structure of a molecule or complex may be described by the set of atoms that best predict the observed diffraction data (that is, which possesses a minimal R value). Files may be created for the structure that defines each atom by its chemical identity, spatial coordinates in three dimensions, root mean squared deviation from the mean observed position and fractional occupancy of the observed position.

Those of skill in the art understand that a set of structure coordinates for an protein, complex or a portion thereof, is a relative set of points that define a shape in three dimensions. Thus, it is possible that an entirely different set of coordinates could define a similar or identical shape. Moreover, slight variations in the individual coordinates may have little affect on overall shape. Such variations in coordinates may be generated because of mathematical manipulations of the structure coordinates. For example, structure coordinates could be manipulated by crystallographic permutations of the structure coordinates, fractionalization of the structure coordinates, integer additions or subtractions to sets of the structure coordinates, inversion of the structure coordinates or any combination of the above. Alternatively, modifications in the crystal structure due to mutations, additions, substitutions, and/or deletions of amino acids, or other changes in any of the components that make up the crystal, could also yield variations in structure coordinates. Such slight variations in the individual coordinates will have little affect on overall shape. If such variations are within an acceptable standard error as compared to the original coordinates, the resulting three-dimensional shape is considered to be structurally equivalent. It should be noted that slight variations in individual structure coordinates of a polypeptide of the invention or a complex thereof would not be expected to significantly alter the nature of modulators that could associate with a druggable region thereof. Thus, for example, a modulator that bound to the active site of a polypeptide of the invention would also be expected to bind to or interfere with another active site whose structure coordinates define a shape that falls within the acceptable error.

A crystal structure of the present invention may be used to make a structural or computer model of the polypeptide, complex or portion thereof. A model may represent the secondary, tertiary and/or quaternary structure of the polypeptide, complex or portion. The configurations of points in space derived from structure coordinates according to the invention can be visualized as, for example, a holographic image, a stereodiagram, a model or a computer-displayed image, and the invention thus includes such images, diagrams or models.

(iii) Structural Equivalents

Various computational analyses can be used to determine whether a molecule or the active site portion thereof is structurally equivalent with respect to its three-dimensional structure, to all or part of a structure of a polypeptide of the invention or a portion thereof.

For the purpose of this invention, any molecule or complex or portion thereof, that has a root mean square deviation of conserved residue backbone atoms (N, Cα, C, O) of less than about 1.75 Å, when superimposed on the relevant backbone atoms described by the reference structure coordinates of a polypeptide of the invention, is considered "structurally equivalent" to the reference molecule. That is to say, the crystal structures of those portions of the two molecules are substantially identical, within acceptable error. Alternatively, the root mean square deviation may be is less than about 1.50, 1.40, 1.25, 1.0, 0.75, 0.5 or 0.35 Å.

The term "root mean square deviation" is understood in the art and means the square root of the arithmetic mean of the squares of the deviations. It is a way to express the deviation or variation from a trend or object.

In another aspect, the present invention provides a scalable three-dimensional configuration of points, at least a portion of said points, and preferably all of said points, derived from structural coordinates of at least a portion of a polypeptide of the invention and having a root mean square deviation from the structure coordinates of the polypeptide of the invention of less than 1.50, 1.40, 1.25, 1.0, 0.75, 0.5 or 0.35 Å. In certain embodiments, the portion of a polypeptide of the invention is 25%, 33%, 50%, 66%, 75%, 85%, 90% or 95% or more of the amino acid residues contained in the polypeptide.

In another aspect, the present invention provides a molecule or complex including a druggable region of a polypeptide of the invention, the druggable region being defined by a set of points having a root mean square deviation of less than about 1.75 Å from the structural coordinates for points representing (a) the backbone atoms of the amino acids contained in a druggable region of a polypeptide of the invention, (b) the side chain atoms (and optionally the Cα atoms) of the amino acids contained in such druggable region, or (c) all the atoms of the amino acids contained in such druggable region. In certain embodiments, only a portion of the amino acids of a druggable region may be included in the set of points, such as 25%, 33%, 50%, 66%, 75%, 85%, 90% or 95% or more of the amino acid residues contained in the druggable region. In certain embodiments, the root mean square deviation may be less than 1.50, 1.40, 1.25, 1.0, 0.75, 0.5, or 0.35 Å. In still other embodiments, instead of a druggable region, a stable domain, fragment or structural motif is used in place of a druggable region.

(iv) Machine Displays and Machine Readable Storage Media

The invention provides a machine-readable storage medium including a data storage material encoded with machine readable data which, when using a machine programmed with instructions for using said data, displays a graphical three-dimensional representation of any of the molecules or complexes, or portions thereof, of this invention. In another embodiment, the graphical three-dimensional representation of such molecule, complex or portion thereof includes the root mean square deviation of certain atoms of such molecule by a specified amount, such as the backbone atoms by less than 0.8 Å. In another embodiment, a structural equivalent of such molecule, complex, or portion thereof, may be displayed. In another embodiment, the portion may include a druggable region of the polypeptide of the invention.

According to one embodiment, the invention provides a computer for determining at least a portion of the structure coordinates corresponding to x-ray diffraction data obtained from a molecule or complex, wherein said computer includes: (a) a machine-readable data storage medium comprising a data storage material encoded with machine-readable data, wherein said data comprises at least a portion of the structural coordinates of a polypeptide of the invention; (b) a machine-readable data storage medium comprising a data storage material encoded with machine-readable data, wherein said data comprises x-ray diffraction data from said molecule or complex; (c) a working memory for storing instructions for processing said machine-readable data of (a) and (b); (d) a central-processing unit coupled to said working memory and to said machine-readable data storage medium of (a) and (b) for performing a Fourier transform of the machine readable data of (a) and for processing said machine readable data of (b) into structure coordinates; and (e) a display coupled to said central-processing unit for displaying said structure coordinates of said molecule or complex. In certain embodiments, the structural coordinates displayed are structurally equivalent to the structural coordinates of a polypeptide of the invention.

In an alternative embodiment, the machine-readable data storage medium includes a data storage material encoded with a first set of machine readable data which includes the Fourier transform of the structure coordinates of a polypeptide of the invention or a portion thereof, and which, when using a machine programmed with instructions for using said data, can be combined with a second set of machine readable data including the x-ray diffraction pattern of a molecule or complex to determine at least a portion of the structure coordinates corresponding to the second set of machine readable data.

For example, a system for reading a data storage medium may include a computer including a central processing unit ("CPU"), a working memory which may be, e.g., RAM (random access memory) or "core" memory, mass storage memory (such as one or more disk drives or CD-ROM drives), one or more display devices (e.g., cathode-ray tube ("CRT") displays, light emitting diode ("LED") displays, liquid crystal displays ("LCDs"), electroluminescent displays, vacuum fluorescent displays, field emission displays ("FEDs"), plasma displays, projection panels, etc.), one or more user input devices (e.g., keyboards, microphones, mice, touch screens, etc.), one or more input lines, and one or more output lines, all of which are interconnected by a conventional bidirectional system bus. The system may be a stand-alone computer, or may be networked (e.g., through local area networks, wide area networks, intranets, extranets, or the internet) to other systems (e.g., computers, hosts, servers, etc.). The system may also include additional computer controlled devices such as consumer electronics and appliances.

Input hardware may be coupled to the computer by input lines and may be implemented in a variety of ways. Machine-readable data of this invention may be inputted via the use of a modem or modems connected by a telephone line or dedicated data line. Alternatively or additionally, the input hardware may include CD-ROM drives or disk drives. In conjunction with a display terminal, a keyboard may also be used as an input device.

Output hardware may be coupled to the computer by output lines and may similarly be implemented by conventional devices. By way of example, the output hardware may include a display device for displaying a graphical representation of an active site of this invention using a program such as QUANTA as described herein. Output hardware might also include a printer, so that hard copy output may be produced, or a disk drive, to store system output for later use.

In operation, a CPU coordinates the use of the various input and output devices, coordinates data accesses from mass storage devices, accesses to and from working memory, and determines the sequence of data processing steps. A number of programs may be used to process the machine-readable data of this invention. Such programs are discussed in reference to the computational methods of drug discovery as described herein. References to components of the hardware system are included as appropriate throughout the following description of the data storage medium.

Machine-readable storage devices useful in the present invention include, but are not limited to, magnetic devices, electrical devices, optical devices, and combinations thereof. Examples of such data storage devices include, but are not limited to, hard disk devices, CD devices, digital video disk devices, floppy disk devices, removable hard disk devices, magneto-optic disk devices, magnetic tape devices, flash memory devices, bubble memory devices, holographic storage devices, and any other mass storage peripheral device. It should be understood that these storage devices include necessary hardware (e.g., drives, controllers, power supplies, etc.) as well as any necessary media (e.g., disks, flash cards, etc.) to enable the storage of data.

In one embodiment, the present invention contemplates a computer readable storage medium comprising structural data, wherein the data include the identity and three-dimensional coordinates of a polypeptide of the invention or portion thereof. In another aspect, the present invention contemplates a database comprising the identity and three-dimensional coordinates of a polypeptide of the invention or a portion thereof. Alternatively, the present invention contemplates a database comprising a portion or all of the atomic coordinates of a polypeptide of the invention or portion thereof.

(v) Structurally Similar Molecules and Complexes

Structural coordinates for a polypeptide of the invention can be used to aid in obtaining structural information about another molecule or complex. This method of the invention allows determination of at least a portion of the three-dimensional structure of molecules or molecular complexes which contain one or more structural features that are similar to structural features of a polypeptide of the invention. Similar structural features can include, for example, regions of amino acid identity, conserved active site or binding site motifs, and similarly arranged secondary structural elements (e.g., α helices and β sheets). Many of the methods described above for determining the structure of a polypeptide of the invention may be used for this purpose as well.

For the present invention, a "structural homolog" is a polypeptide that contains one or more amino acid substitutions, deletions, additions, or rearrangements with respect to the amino acid sequence of SEQ ID NO: 4 or other polypeptide of the invention, but that, when folded into its native conformation, exhibits or is reasonably expected to exhibit at least a portion of the tertiary (three-dimensional) structure of the polypeptide encoded by SEQ ID NO: 4 or such other polypeptide of the invention. For example, structurally homologous molecules can contain deletions or additions of one or more contiguous or noncontiguous

amino acids, such as a loop or a domain. Structurally homologous molecules also include modified polypeptide molecules that have been chemically or enzymatically derivatized at one or more constituent amino acids, including side chain modifications, backbone modifications, and N- and C-terminal modifications including acetylation, hydroxylation, methylation, amidation, and the attachment of carbohydrate or lipid moieties, cofactors, and the like.

By using molecular replacement, all or part of the structure coordinates of a polypeptide of the invention can be used to determine the structure of a crystallized molecule or complex whose structure is unknown more quickly and efficiently than attempting to determine such information *ab initio*. For example, in one embodiment this invention provides a method of utilizing molecular replacement to obtain structural information about a molecule or complex whose structure is unknown including: (a) crystallizing the molecule or complex of unknown structure; (b) generating an x-ray diffraction pattern from said crystallized molecule or complex; and (c) applying at least a portion of the structure coordinates for a polypeptide of the invention to the x-ray diffraction pattern to generate a three-dimensional electron density map of the molecule or complex whose structure is unknown.

In another aspect, the present invention provides a method for generating a preliminary model of a molecule or complex whose structure coordinates are unknown, by orienting and positioning the relevant portion of a polypeptide of the invention within the unit cell of the crystal of the unknown molecule or complex so as best to account for the observed x-ray diffraction pattern of the crystal of the molecule or complex whose structure is unknown.

Structural information about a portion of any crystallized molecule or complex that is sufficiently structurally similar to a portion of a polypeptide of the invention may be resolved by this method. In addition to a molecule that shares one or more structural features with a polypeptide of the invention, a molecule that has similar bioactivity, such as the same catalytic activity, substrate specificity or ligand binding activity as a polypeptide of the invention, may also be sufficiently structurally similar to a polypeptide of the invention to permit use of the structure coordinates for a polypeptide of the invention to solve its crystal structure.

In another aspect, the method of molecular replacement is utilized to obtain structural information about a complex containing a polypeptide of the invention, such as a

complex between a modulator and a polypeptide of the invention (or a domain, fragment, ortholog, homolog etc. thereof). In certain instances, the complex includes a polypeptide of the invention (or a domain, fragment, ortholog, homolog etc. thereof) co-complexed with a modulator. For example, in one embodiment, the present invention contemplates a method for making a crystallized complex comprising a polypeptide of the invention, or a fragment thereof, and a compound having a molecular weight of less than 5 kDa, the method comprising: (a) crystallizing a polypeptide of the invention such that the crystals will diffract x-rays to a resolution of 3.5 Å or better; and (b) soaking the crystal in a solution comprising the compound having a molecular weight of less than 5 kDa, thereby producing a crystallized complex comprising the polypeptide and the compound.

Using homology modeling, a computer model of a structural homolog or other polypeptide can be built or refined without crystallizing the molecule. For example, in another aspect, the present invention provides a computer-assisted method for homology modeling a structural homolog of a polypeptide of the invention including: aligning the amino acid sequence of a known or suspected structural homolog with the amino acid sequence of a polypeptide of the invention and incorporating the sequence of the homolog into a model of a polypeptide of the invention derived from atomic structure coordinates to yield a preliminary model of the homolog; subjecting the preliminary model to energy minimization to yield an energy minimized model; remodeling regions of the energy minimized model where stereochemistry restraints are violated to yield a final model of the homolog.

In another embodiment, the present invention contemplates a method for determining the crystal structure of a homolog of a polypeptide having SEQ ID NO: 2 or SEQ ID NO: 4, or equivalent thereof, the method comprising: (a) providing the three dimensional structure of a crystallized polypeptide having SEQ ID NO: 2 or SEQ ID NO: 4, or a fragment thereof; (b) obtaining crystals of a homologous polypeptide comprising an amino acid sequence that is at least 80% identical to the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4 such that the three dimensional structure of the crystallized homologous polypeptide may be determined to a resolution of 3.5 Å or better; and (c) determining the three dimensional structure of the crystallized homologous polypeptide by x-ray crystallography based on the atomic coordinates of the three dimensional structure provided in step (a). In certain instances of the foregoing method, the atomic coordinates for the homologous polypeptide have a root mean square deviation from

the backbone atoms of the polypeptide having SEQ ID NO: 2 or SEQ ID NO: 4, or a fragment thereof, of not more than 1.5 Å for all backbone atoms shared in common with the homologous polypeptide and the polypeptide having SEQ ID NO: 2 or SEQ ID NO: 4, or a fragment thereof.

(vi) NMR Analysis Using X-ray Structural Data

In another aspect, the structural coordinates of a known crystal structure may be applied to nuclear magnetic resonance data to determine the three dimensional structures of polypeptides with uncharacterized or incompletely characterized structure. (See for example, Wuthrich, 1986, John Wiley and Sons, New York: 176-199; Pflugrath et al., 1986, J. Molecular Biology 189: 383-386; Kline et al., 1986 J. Molecular Biology 189:377-382). While the secondary structure of a polypeptide may often be determined by NMR data, the spatial connections between individual pieces of secondary structure are not as readily determined. The structural coordinates of a polypeptide defined by x-ray crystallography can guide the NMR spectroscopist to an understanding of the spatial interactions between secondary structural elements in a polypeptide of related structure. Information on spatial interactions between secondary structural elements can greatly simplify NOE data from two-dimensional NMR experiments. In addition, applying the structural coordinates after the determination of secondary structure by NMR techniques simplifies the assignment of NOE's relating to particular amino acids in the polypeptide sequence.

In an embodiment, the invention relates to a method of determining three dimensional structures of polypeptides with unknown structures, by applying the structural coordinates of a crystal of the present invention to nuclear magnetic resonance data of the unknown structure. This method comprises the steps of: (a) determining the secondary structure of an unknown structure using NMR data; and (b) simplifying the assignment of through-space interactions of amino acids. The term "through-space interactions" defines the orientation of the secondary structural elements in the three dimensional structure and the distances between amino acids from different portions of the amino acid sequence. The term "assignment" defines a method of analyzing NMR data and identifying which amino acids give rise to signals in the NMR spectrum.

For all of this section on x-ray cystallography, see also Brooks et al. (1983) *J Comput Chem* 4:187-217; Weiner et al (1981) *J. Comput. Chem.* 106: 765; Eisenfield et al. (1991) *Am J Physiol* 261:C376-386; Lybrand (1991) *J Pharm Belg* 46:49-54; Froimowitz (1990) *Biotechniques* 8:640-644; Burbam et al. (1990) *Proteins* 7:99-111; Pedersen (1985)

Environ Health Perspect 61:185-190; and Kini et al. (1991) *J Biomol Struct Dyn* 9:475-488; Ryckaert et al. (1977) *J Comput Phys* 23:327; Van Gunsteren et al. (1977) *Mol Phys* 34:1311; Anderson (1983) *J Comput Phys* 52:24; J. Mol. Biol. 48: 442-453, 1970; Dayhoff et al., Meth. Enzymol. 91: 524-545, 1983; Henikoff and Henikoff, Proc. Nat. Acad. Sci. USA 89: 10915-10919, 1992; J. Mol. Biol. 233: 716-738, 1993; Methods in Enzymology, Volume 276, Macromolecular crystallography, Part A, ISBN 0-12-182177-3 and Volume 277, Macromolecular crystallography, Part B, ISBN 0-12-182178-1, Eds. Charles W. Carter, Jr. and Robert M. Sweet (1997), Academic Press, San Diego; Pfuetzner, et al., J. Biol. Chem. 272: 430-434 (1997).

6. Interacting Proteins

The present invention also provides methods for isolating specific protein interactors of a polypeptide of the invention, and complexes comprising a polypeptide of the invention and one or more interacting proteins. In one aspect, the present invention contemplates an isolated protein complex comprising a polypeptide of the invention and at least one protein that interacts with the polypeptide of the invention. The protein may be naturally-occurring. The interacting protein may be of *P. aeruginosa* origin. Alternatively, the interacting protein may be of mammalian origin or human origin. Either the polypeptide of the invention or the interacting protein or both may be a fusion protein.

The present invention contemplates a method for identifying a protein capable of interacting with a polypeptide of the invention or a fragment thereof, the method comprising: (a) exposing a sample to a solid substrate coupled to a polypeptide of the invention or a fragment thereof under conditions which promote protein-protein interactions; (b) washing the solid substrate so as to remove any polypeptides interacting non-specifically with the polypeptide or fragment; (c) eluting the polypeptides which specifically interact with the polypeptide or fragment; and (d) identifying the interacting protein. The sample may be an extract of *P. aeruginosa*, a mammalian cell extract, a human cell extract, a purified protein (or a fragment thereof), or a mixture of purified proteins (or fragments thereof). The interacting protein may be identified by a number of methods, including mass spectrometry or protein sequencing.

In another aspect, the present invention contemplates a method for identifying a protein capable of interacting with a polypeptide of present invention or a fragment thereof, the method comprising: (a) subjecting a sample to protein-affinity chromatography on

multiple columns, the columns having a polypeptide of the invention or a fragment thereof coupled to the column matrix in varying concentrations, and eluting bound components of the extract from the columns; (b) separating the components to isolate a polypeptide capable of interacting with the polypeptide or fragment; and (c) analyzing the interacting protein by mass spectrometry to identify the interacting protein. In certain instances, the foregoing method will use polyacrylamide gel electrophoresis without SDS.

In another aspect, the present invention contemplates a method for identifying a protein capable of interacting with a polypeptide of the invention, the method comprising: (a) subjecting a cellular extract or extracellular fluid to protein-affinity chromatography on multiple columns, the columns having a polypeptide of the invention or a fragment thereof coupled to the column matrix in varying concentrations, and eluting bound components of the extract from the columns; (b) gel-separating the components to isolate an interacting protein; wherein the interacting protein is observed to vary in amount in direct relation to the concentration of coupled polypeptide or fragment; (c) digesting the interacting protein to give corresponding peptides; (d) analyzing the peptides by MALDI-TOF mass spectrometry or post source decay to determine the peptide masses; and (d) performing correlative database searches with the peptide, or peptide fragment, masses, whereby the interacting protein is identified based on the masses of the peptides or peptide fragments. The foregoing method may include the further step of including the identifies of any interacting proteins into a relational database.

In another aspect, the invention further contemplates a method for identifying modulators of a protein complex, the method comprising: (a) contacting a protein complex comprising a polypeptide of the invention and an interacting protein with one or more test compounds; and (b) determining the effect of the test compound on (i) the activity of the protein complex, (ii) the amount of the protein complex, (iii) the stability of the protein complex, (iv) the conformation of the protein complex, (v) the activity of at least one polypeptide included in the protein complex, (vi) the conformation of at least one polypeptide included in the protein complex, (vii) the intracellular localization of the protein complex or a component thereof, (viii) the transcription level of a gene dependent on the complex, and/or (ix) the level of second messenger levels in a cell; thereby identifying modulators of the protein complex. The foregoing method may be carried out *in vitro* or *in vivo* as appropriate.

Typically, it will be desirable to immobilize a polypeptide of the invention to facilitate separation of complexes comprising a polypeptide of the invention from uncomplexed forms of the interacting proteins, as well as to accommodate automation of the assay. The polypeptide of the invention, or ligand, may be immobilized onto a solid support (e.g., column matrix, microtiter plate, slide, etc.). In certain embodiments, the ligand may be purified. In certain instances, a fusion protein may be provided which adds a domain that permits the ligand to be bound to a support.

In various *in vitro* embodiments, the set of proteins engaged in a protein-protein interaction comprises a cell extract, a clarified cell extract, or a reconstituted protein mixture of at least semi-purified proteins. By semi-purified, it is meant that the proteins utilized in the reconstituted mixture have been previously separated from other cellular or viral proteins. For instance, in contrast to cell lysates, the proteins involved in a protein-protein interaction are present in the mixture to at least about 50% purity relative to all other proteins in the mixture, and more preferably are present in greater, even 90-95%, purity. In certain embodiments of the subject method, the reconstituted protein mixture is derived by mixing highly purified proteins such that the reconstituted mixture substantially lacks other proteins (such as of cellular or viral origin) which might interfere with or otherwise alter the ability to measure activity resulting from the given protein-protein interaction.

Complex formation involving a polypeptide of the invention and another component polypeptide or a substrate polypeptide, may be detected by a variety of techniques. For instance, modulation in the formation of complexes can be quantitated using, for example, detectably labeled proteins (e.g. radiolabeled, fluorescently labeled, or enzymatically labeled), by immunoassay, or by chromatographic detection.

The present invention also provides assays for identifying molecules which are modulators of a protein-protein interaction involving a polypeptide of the invention, or are a modulator of the role of the complex comprising a polypeptide of the invention in the infectivity or pathogenicity of *P. aeruginosa*. In one embodiment, the assay detects agents which inhibit formation or stabilization of a protein complex comprising a polypeptide of the invention and one or more additional proteins. In another embodiment, the assay detects agents which modulate the intrinsic biological activity of a protein complex comprising a polypeptide of the invention, such as an enzymatic activity, binding to other cellular components, cellular compartmentalization, signal transduction, and the like. Such

modulators may be used, for example, in the treatment of *P. aeruginosa* related diseases or disorders. In certain embodiments, the compound is a mechanism based inhibitor which chemically alters one member of a protein-protein interaction involving a polypeptide of the invention and which is a specific inhibitor of that member, e.g. has an inhibition constant about 10-fold, 100-fold, or 1000-fold different compared to homologous proteins.

In one embodiment, proteins that interact with a polypeptide of the invention may be isolated using immunoprecipitation. A polypeptide of the invention may be expressed in *P. aeruginosa*, or in a heterologous system. The cells expressing a polypeptide of the invention are then lysed under conditions which maintain protein-protein interactions, and complexes comprising a polypeptide of the invention are isolated. For example, a polypeptide of the invention may be expressed in mammalian cells, including human cells, in order to identify mammalian proteins that interact with a polypeptide of the invention and therefore may play a role in *P. aeruginosa* infectivity or proliferation. In one embodiment, a polypeptide of the invention is expressed in the cell type for which it is desirable to find interacting proteins. For example, a polypeptide of the invention may be expressed in *P. aeruginosa* in order to find *P. aeruginosa* derived interacting proteins.

In an alternative embodiment, a polypeptide of the invention is expressed and purified and then mixed with a potential interacting protein or mixture of proteins to identify complex formation. The potential interacting protein may be a single purified or semi-purified protein, or a mixture of proteins, including a mixture of purified or semi-purified proteins, a cell lysate, a clarified cell lysate, a semi-purified cell lysate, etc.

In certain embodiments, it may be desirable to use a tagged version of a polypeptide of the invention in order to facilitate isolation of complexes from the reaction mixture. Suitable tags for immunoprecipitation experiments include HA, myc, FLAG, HIS, GST, protein A, protein G, etc. Immunoprecipitation from a cell lysate or other protein mixture may be carried out using an antibody specific for a polypeptide of the invention or using an antibody which recognizes a tag to which a polypeptide of the invention is fused (e.g., anti-HA, anti-myc, anti-FLAG, etc.). Antibodies specific for a variety of tags are known to the skilled artisan and are commercially available from a number of sources. In the case where a polypeptide of the invention is fused to a His, GST, or protein A/G tag, immunoprecipitation may be carried out using the appropriate affinity resin (e.g., beads functionalized with Ni, glutathione, Fc region of IgG, etc.). Test compounds which modulate a protein-protein interaction involving a polypeptide of the invention may be

identified by carrying out the immunoprecipitation reaction in the presence and absence of the test agent and comparing the level and/or activity of the protein complex between the two reactions.

In another embodiment, proteins that interact with a polypeptide of the invention may be identified using affinity chromatography. Some examples of such chromatography are described in USSN 09/727,812, filed November 30, 2000, and the PCT Application filed November 30, 2001 and entitled "Methods for Systematic Identification of Protein-Protein Interactions and other Properties", which claims priority to such U.S. application.

In one aspect, for affinity chromatography using a solid support, a polypeptide of the invention or a fragment thereof may be attached by a variety of means known to those of skill in the art. For example, the polypeptide may be coupled directly (through a covalent linkage) to commercially available pre-activated resins as described in Formosa et al., Methods in Enzymology 1991, 208, 24-45; Sopta et al, J. Biol. Chem. 1985, 260, 10353-60; Archambault et al., Proc. Natl. Acad. Sci. USA 1997, 94, 14300-5. Alternatively, the polypeptide may be tethered to the solid support through high affinity binding interactions. If the polypeptide is expressed fused to a tag, such as GST, the fusion tag can be used to anchor the polypeptide to the matrix support, for example Sepharose beads containing immobilized glutathione. Solid supports that take advantage of these tags are commercially available.

In another aspect, the support to which a polypeptide may be immobilized is a soluble support, which may facilitate certain steps performed in the methods of the present invention. For example, the soluble support may be soluble in the conditions employed to create a binding interaction between a target and the polypeptide, and then used under conditions in which it is a solid for elution of the proteins or other biological materials that bind to a polypeptide.

The concentration of the coupled polypeptide may have an affect on the sensitivity of the method. In certain embodiments, to detect interactions most efficiently, the concentration of the polypeptide bound to the matrix should be at least 10-fold higher than the K_d of the interaction. Thus, the concentration of the polypeptide bound to the matrix should be highest for the detection of the weakest protein-protein interactions. However, if the concentration of the immobilized polypeptide is not as high as may be ideal, it may still be possible to observe protein-protein interactions of interest by, for example, increasing the concentration of the polypeptide or other moiety that interacts with the coupled

polypeptide. The level of detection will of course vary with each different polypeptide, interactor, conditions of the assay, etc. In certain instances, the interacting protein binds to the polypeptide with a K_d of about 10^{-5} M to about 10^{-8} M or 10^{-10} M.

In another aspect, the coupling may be done at various ratios of the polypeptide to the resin. An upper limit of the protein : resin ratio may be determined by the isoelectric point and the ionic nature of the protein, although it may be possible to achieve higher polypeptide concentrations by use of various methods.

In certain embodiments, several concentrations of the polypeptide immobilized on a solid or soluble support may be used. One advantage of using multiple concentrations, although not a requirement, is that one may be able to obtain an estimate for the strength of the protein-protein interaction that is observed in the affinity chromatography experiment. Another advantage of using multiple concentrations is that a binding curve which has the proper shape may indicate that the interaction that is observed is biologically important rather than a spurious interaction with denatured protein.

In one example of such an embodiment, a series of columns may be prepared with varying concentrations of polypeptide (mg polypeptide/ml resin volume). The number of columns employed may be between 2 to 8, 10, 12, 15, 25 or more, each with a different concentration of attached polypeptide. Larger numbers of columns may be used if appropriate for the polypeptide being examined, and multiple columns may be used with the same concentration as any methods may require. In certain embodiments, 4 to 6 columns are prepared with varying concentrations of polypeptide. In another aspect of this embodiment, two control columns may be prepared: one that contains no polypeptide and a second that contains the highest concentration of polypeptide but is not treated with extract. After elution of the columns and separation of the eluent components (by one of the methods described below), it may be possible to distinguish the interacting proteins (if any) from the non-specific bound proteins as follows. The concentration of the interacting proteins, as determined by the intensity of the band on the gel, will increase proportionally to the increase in polypeptide concentration but will be missing from the second control column. This allows for the identification of unknown interacting proteins.

The method of the invention may be used for small-scale analysis. A variety of column sizes, types, and geometries may be used. In addition, other vessel shapes and sizes having a smaller scale than is usually found in laboratory experiments may be used as well, including a plurality of wells in a plate. For high throughput analysis, it is advantageous to

use small volumes, from about 20, 30, 50, 80 or 100 µl. Larger or small volumes may be used, as necessary, and it may be possible to achieve high throughput analysis using them. The entire affinity chromatography procedure may be automated by assembling the micro-columns into an array (e.g. with 96 micro-column arrays).

A variety of materials may be used as the source of potential interacting proteins. In one embodiment, a cellular extract or extracellular fluid may be used. The choice of starting material for the extract may be based upon the cell or tissue type or type of fluid that would be expected to contain proteins that interact with the target protein. Micro-organisms or other organisms are grown in a medium that is appropriate for that organism and can be grown in specific conditions to promote the expression of proteins that may interact with the target protein. Exemplary starting material that may be used to make a suitable extract are: 1) one or more types of tissue derived from an animal, plant, or other multi-cellular organism, 2) cells grown in tissue culture that were derived from an animal or human, plant or other source, 3) micro-organisms grown in suspension or non-suspension cultures, 4) virus-infected cells, 5) purified organelles (including, but not restricted to nuclei, mitochondria, membranes, Golgi, endoplasmic reticulum, lysosomes, or peroxisomes) prepared by differential centrifugation or another procedure from animal, plant or other kinds of eukaryotic cells, 6) serum or other bodily fluids including, but not limited to, blood, urine, semen, synovial fluid, cerebrospinal fluid, amniotic fluid, lymphatic fluid or interstitial fluid. In other embodiments, a total cell extract may not be the optimal source of interacting proteins. For example, if the ligand is known to act in the nucleus, a nuclear extract can provide a 10-fold enrichment of proteins that are likely to interact with the ligand. In addition, proteins that are present in the extract in low concentrations may be enriched using another chromatographic method to fractionate the extract before screening various pools for an interacting protein.

Extracts are prepared by methods known to those of skill in the art. The extracts may be prepared at a low temperature (e.g., 4°C) in order to retard denaturation or degradation of proteins in the extract. The pH of the extract may be adjusted to be appropriate for the body fluid or tissue, cellular, or organellar source that is used for the procedure (e.g. pH 7-8 for cytosolic extracts from mammals, but low pH for lysosomal extracts). The concentration of chaotropic or non-chaotropic salts in the extracting solution may be adjusted so as to extract the appropriate sets of proteins for the procedure. Glycerol may be added to the extract, as it aids in maintaining the stability of many proteins and also

reduces background non-specific binding. Both the lysis buffer and column buffer may contain protease inhibitors to minimize proteolytic degradation of proteins in the extract and to protect the polypeptide. Appropriate co-factors that could potentially interact with the interacting proteins may be added to the extracting solution. One or more nucleases or another reagent may be added to the extract, if appropriate, to prevent protein-protein interactions that are mediated by nucleic acids. Appropriate detergents or other agents may be added to the solution, if desired, to extract membrane proteins from the cells or tissue. A reducing agent (e.g. dithiothreitol or 2-mercaptoethanol or glutathione or other agent) may be added. Trace metals or a chelating agent may be added, if desired, to the extracting solution.

Usually, the extract is centrifuged in a centrifuge or ultracentrifuge or filtered to provide a clarified supernatant solution. This supernatant solution may be dialyzed using dialysis tubing, or another kind of device that is standard in the art, against a solution that is similar to, but may not be identical with, the solution that was used to make the extract. The extract is clarified by centrifugation or filtration again immediately prior to its use in affinity chromatography.

In some cases, the crude lysate will contain small molecules that can interfere with the affinity chromatography. This can be remedied by precipitating proteins with ammonium sulfate, centrifugation of the precipitate, and re-suspending the proteins in the affinity column buffer followed by dialysis. An additional centrifugation of the sample may be needed to remove any particulate matter prior to application to the affinity columns.

The amount of cell extract applied to the column may be important for any embodiment. If too little extract is applied to the column and the interacting protein is present at low concentration, the level of interacting protein retained by the column may be difficult to detect. Conversely, if too much extract is applied to the column, protein may precipitate on the column or competition by abundant interacting proteins for the limited amount of protein ligand may result in a difficulty in detecting minor species.

The columns functionalized with a polypeptide of the invention are loaded with protein extract from an appropriate source that has been dialyzed against a buffer that is consistent with the nature of the expected interaction. The pH, salt concentrations and the presence or absence of reducing and chelating agents, trace metals, detergents, and co-factors may be adjusted according to the nature of the expected interaction. Most commonly, the pH and the ionic strength are chosen so as to be close to physiological for

the source of the extract. The extract is most commonly loaded under gravity onto the columns at a flow rate of about 4-6 column volumes per hour, but this flow rate can be adjusted for particular circumstances in an automated procedure.

The volume of the extract that is loaded on the columns can be varied but is most commonly equivalent to about 5 to 10 column volumes. When large volumes of extract are loaded on the columns, there is often an improvement in the signal-to-noise ratio because more protein from the extract is available to bind to the protein ligand, whereas the background binding of proteins from the extract to the solid support saturates with low amounts of extract.

A control column may be included that contains the highest concentration of protein ligand, but buffer rather than extract is loaded onto this column. The elutions (eluates) from this column will contain polypeptide that failed to be attached to the column in a covalent manner, but no proteins that are derived from the extract.

The columns may be washed with a buffer appropriate to the nature of the interaction being analyzed, usually, but not necessarily, the same as the loading buffer. An elution buffer with an appropriate pH, glycerol, and the presence or absence of reducing agent, chelating agent, cofactors, and detergents are all important considerations. The columns may be washed with anywhere from about 5 to 20 column volumes of each wash buffer to eliminate unbound proteins from the natural extract. The flow rate of the wash is usually adjusted to about 4 to 6 column volumes per hour by using gravity or an automated procedure, but other flow rates are possible in specific circumstances.

In order to elute the proteins that have been retained by the column, the interactions between the extract proteins and the column ligand should be disrupted. This is performed by eluting the column with a solution of salt or detergent. Retention of activity by the eluted proteins may require the presence of glycerol and a buffer of appropriate pH, as well as proper choices of ionic strength and the presence or absence of appropriate reducing agent, chelating agent, trace metals, cofactors, detergents, chaotropic agents, and other reagents. If physical identification of the bound proteins is the objective, the elution may be performed sequentially, first with buffer of high ionic strength and then with buffer containing a protein denaturant, most commonly, but not restricted to sodium dodecyl sulfate (SDS), urea, or guanidine hydrochloride. In certain instances, the column is eluted with a protein denaturant, particularly SDS, for example as a 1% SDS solution. Using only the SDS wash, and omitting the salt wash, may result in SDS-gels that have higher

resolution (sharper bands with less smearing). Also, using only the SDS wash results in half as many samples to analyze. The volume of the eluting solution may be varied but is normally about 2 to 4 column volumes. For 20 ml columns, the flow rate of the eluting procedures are most commonly about 4 to 6 column volumes per hour, under gravity, but can be varied in an automated procedure.

The proteins from the extract that were bound to and are eluted from the affinity columns may be most easily resolved for identification by an electrophoresis procedure, but this procedure may be modified, replaced by another suitable method, or omitted. Any of the denaturing or non-denaturing electrophoresis procedures that are standard in the art may be used for this purpose, including SDS-PAGE, gradient gels, capillary electrophoresis, and two-dimensional gels with isoelectric focusing in the first dimension and SDS-PAGE in the second. Typically, the individual components in the column eluent are separated by polyacrylamide gel electrophoresis.

After electrophoresis, protein bands or spots may be visualized using any number of methods know to those of skill in the art, including staining techniques such as Coomassie blue or silver staining, or some other agent that is standard in the art. Alternatively, autoradiography can be used for visualizing proteins isolated from organisms cultured on media containing a radioactive label, for example $^{35}SO_4^{2-}$ or ^{35}[S]methionine, that is incorporated into the proteins. The use of radioactively labeled extract allows a distinction to be made between extract proteins that were retained by the column and proteolytic fragments of the ligand that may be released from the column.

Protein bands that are derived from the extract (i.e. it did not elute from the control column that was not loaded with protein from the extract) and bound to an experimental column that contained polypeptide covalently attached to the solid support, and did not bind to a control column that did not contain any polypeptide, may be excised from the stained electrophoretic gel and further characterized.

To identify the protein interactor by mass spectrometry, it may be desirable to reduce the disulfide bonds of the protein followed by alkylation of the free thiols prior to digestion of the protein with protease. The reduction may be performed by treatment of the gel slice with a reducing agent, for example with dithiothreitol, whereupon, the protein is alkylated by treating the gel slice with a suitable alkylating agent, for example iodoacetamide.

Prior to analysis by mass spectrometry, the protein may be chemically or enzymatically digested. The protein sample in the gel slice may be subjected to *in-gel* digestion. Shevchenko A. et al., Mass Spectrometric Sequencing of Proteins from Silver Stained Polyacrylamide Gels. Analytical Chemistry 1996, 58, 850-858. One method of digestion is by treatment with the enzyme trypsin. The resulting peptides are extracted from the gel slice into a buffer.

The peptide fragments may be purified, for example by use of chromatography. A solid support that differentially binds the peptides and not the other compounds derived from the gel slice, the protease reaction or the peptide extract may be used. The peptides may be eluted from the solid support into a small volume of a solution that is compatible with mass spectrometry (e.g. 50% acetonitrile/0.1% trifluoroacetic acid).

The preparation of a protein sample from a gel slice that is suitable for mass spectrometry may also be done by an automated procedure.

Peptide samples derived from gel slices may be analyzed by any one of a variety of techniques in mass spectrometry as further described above. This technique may be used to assign function to an unknown protein based upon the known function of the interacting protein in the same or a homologous/orthologous organism.

Eluates from the affinity chromatography columns may also be analyzed directly without resolution by electrophoretic methods, by proteolytic digestion with a protease in solution, followed by applying the proteolytic digestion products to a reverse phase column and eluting the peptides from the column.

In yet another embodiment, proteins that interact with a polypeptide of the invention may be identified using an interaction trap assay (see also, U.S. Patent NO: 5,283,317; Zervos *et al.* (1993) *Cell* 72:223-232; Madura *et al.* (1993) *J Biol Chem* 268:12046-12054; Bartel *et al.* (1993) *Biotechniques* 14:920-924; and Iwabuchi *et al.* (1993) *Oncogene* 8:1693-1696).

In another embodiment, a method of the present invention makes use of chimeric genes which express hybrid proteins. To illustrate, a first hybrid gene comprises the coding sequence for a DNA-binding domain of a transcriptional activator fused in frame to the coding sequence for a "bait" protein, e.g., a polypeptide of the invention of sufficient length to bind to a potential interacting protein. The second hybrid protein encodes a transcriptional activation domain fused in frame to a gene encoding a "fish" protein, e.g., a potential interacting protein of sufficient length to interact with a polypeptide of the

invention portion of the bait fusion protein. If the bait and fish proteins are able to interact, e.g., form a protein-protein interaction, they bring into close proximity the two domains of the transcriptional activator. This proximity causes transcription of a reporter gene which is operably linked to a transcriptional regulatory site responsive to the transcriptional activator, and expression of the reporter gene can be detected and used to score for the interaction of the bait and fish proteins.

In accordance with the present invention, the method includes providing a host cell, typically a yeast cell, e.g., *Kluyverei lactis, Schizosaccharomyces pombe, Ustilago maydis, Saccharomyces cerevisiae, Neurospora crassa, Aspergillus niger, Aspergillus nidulans, Pichia pastoris, Candida tropicalis, and Hansenula polymorpha*, though most preferably *S cerevisiae* or *S. pombe*. The host cell contains a reporter gene having a binding site for the DNA-binding domain of a transcriptional activator used in the bait protein, such that the reporter gene expresses a detectable gene product when the gene is transcriptionally activated. The first chimeric gene may be present in a chromosome of the host cell, or as part of an expression vector.

The host cell also contains a first chimeric gene which is capable of being expressed in the host cell. The gene encodes a chimeric protein, which comprises (a) a DNA-binding domain that recognizes the responsive element on the reporter gene in the host cell, and (b) a bait protein (e.g., a polypeptide of the invention).

A second chimeric gene is also provided which is capable of being expressed in the host cell, and encodes the "fish" fusion protein. In one embodiment, both the first and the second chimeric genes are introduced into the host cell in the form of plasmids. Preferably, however, the first chimeric gene is present in a chromosome of the host cell and the second chimeric gene is introduced into the host cell as part of a plasmid.

The DNA-binding domain of the first hybrid protein and the transcriptional activation domain of the second hybrid protein may be derived from transcriptional activators having separable DNA-binding and transcriptional activation domains. For instance, these separate DNA-binding and transcriptional activation domains are known to be found in the yeast GAL4 protein, and are known to be found in the yeast GCN4 and ADR1 proteins. Many other proteins involved in transcription also have separable binding and transcriptional activation domains which make them useful for the present invention, and include, for example, the LexA and VP16 proteins. It will be understood that other (substantially) transcriptionally-inert DNA-binding domains may be used in the subject

constructs; such as domains of ACE1, λcI, lac repressor, jun or fos. In another embodiment, the DNA-binding domain and the transcriptional activation domain may be from different proteins. The use of a LexA DNA binding domain provides certain advantages. For example, in yeast, the LexA moiety contains no activation function and has no known affect on transcription of yeast genes. In addition, use of LexA allows control over the sensitivity of the assay to the level of interaction (see, for example, the Brent *et al.* PCT publication WO94/10300).

In certain embodiments, any enzymatic activity associated with the bait or fish proteins is inactivated, e.g., dominant negative or other mutants of a protein-protein interaction component can be used.

Continuing with the illustrative example, a polypeptide of the invention-mediated interaction, if any, between the bait and fish fusion proteins in the host cell, causes the activation domain to activate transcription of the reporter gene. The method is carried out by introducing the first chimeric gene and the second chimeric gene into the host cell, and subjecting that cell to conditions under which the bait and fish fusion proteins and are expressed in sufficient quantity for the reporter gene to be activated. The formation of a protein complex containing a polypeptide of the invention results in a detectable signal produced by the expression of the reporter gene.

In still further embodiments, the protein-protein interaction of interest is generated in whole cells, taking advantage of cell culture techniques to support the subject assay. For example, the protein-protein interaction of interest can be constituted in a prokaryotic or eukaryotic cell culture system. Advantages to generating the protein complex in an intact cell includes the ability to screen for inhibitors of the level or activity of the complex which are functional in an environment more closely approximating that which therapeutic use of the inhibitor would require, including the ability of the agent to gain entry into the cell. Furthermore, certain of the *in vivo* embodiments of the assay are amenable to high through-put analysis of candidate agents.

The components of the protein complex comprising a polypeptide of the invention can be endogenous to the cell selected to support the assay. Alternatively, some or all of the components can be derived from exogenous sources. For instance, fusion proteins can be introduced into the cell by recombinant techniques (such as through the use of an expression vector), as well as by microinjecting the fusion protein itself or mRNA encoding the fusion protein. Moreover, in the whole cell embodiments of the subject assay, the

reporter gene construct can provide, upon expression, a selectable marker. Such embodiments of the subject assay are particularly amenable to high through-put analysis in that proliferation of the cell can provide a simple measure of the protein-protein interaction.

The amount of transcription from the reporter gene may be measured using any method known to those of skill in the art to be suitable. For example, specific mRNA expression may be detected using Northern blots or specific protein product may be identified by a characteristic stain, western blots or an intrinsic activity. In certain embodiments, the product of the reporter gene is detected by an intrinsic activity associated with that product. For instance, the reporter gene may encode a gene product that, by enzymatic activity, gives rise to a detection signal based on color, fluorescence, or luminescence.

The interaction trap assay of the invention may also be used to identify test agents capable of modulating formation of a complex comprising a polypeptide of the invention. In general, the amount of expression from the reporter gene in the presence of the test compound is compared to the amount of expression in the same cell in the absence of the test compound. Alternatively, the amount of expression from the reporter gene in the presence of the test compound may be compared with the amount of transcription in a substantially identical cell that lacks a component of the protein-protein interaction involving a polypeptide of the invention.

7. Antibodies

Another aspect of the invention pertains to antibodies specifically reactive with a polypeptide of the invention. For example, by using peptides based on a polypeptide of the invention, e.g., having an amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 4 or an immunogenic fragment thereof, antisera or monoclonal antibodies may be made using standard methods. An exemplary immunogenic fragment may contain eight, ten or more consecutive amino acid residues of SEQ ID NO: 2 or SEQ ID NO: 4. Certain fragments that are predicted to be immunogenic for the subject amino acid sequences (predicted) are set forth in Table 2 contained in FIGURE 7

The term "antibody" as used herein is intended to include fragments thereof which are also specifically reactive with a polypeptide of the invention. Antibodies can be fragmented using conventional techniques and the fragments screened for utility in the same manner as is suitable for whole antibodies. For example, $F(ab')_2$ fragments can be

generated by treating antibody with pepsin. The resulting $F(ab')_2$ fragment can be treated to reduce disulfide bridges to produce Fab' fragments. The antibody of the present invention is further intended to include bispecific and chimeric molecules, as well as single chain (scFv) antibodies. Also within the scope of the invention are trimeric antibodies, humanized antibodies, human antibodies, and single chain antibodies. All of these modified forms of antibodies as well as fragments of antibodies are intended to be included in the term "antibody".

In one aspect, the present invention contemplates a purified antibody that binds specifically to a polypeptide of the invention and which does not substantially cross-react with a protein which is less than about 80%, or less than about 90%, identical to SEQ ID NO: 2 or SEQ ID NO: 4. In another aspect, the present invention contemplates an array comprising a substrate having a plurality of address, wherein at least one of the addresses has disposed thereon a purified antibody that binds specifically to a polypeptide of the invention.

Antibodies may be elicited by methods known in the art. For example, a mammal such as a mouse, a hamster or rabbit may be immunized with an immunogenic form of a polypeptide of the invention (e.g., an antigenic fragment which is capable of eliciting an antibody response). Alternatively, immunization may occur by using a nucleic acid of the acid, which presumably *in vivo* expresses the polypeptide of the invention giving rise to the immunogenic response observed. Techniques for conferring immunogenicity on a protein or peptide include conjugation to carriers or other techniques well known in the art. For instance, a peptidyl portion of a polypeptide of the invention may be administered in the presence of adjuvant. The progress of immunization may be monitored by detection of antibody titers in plasma or serum. Standard ELISA or other immunoassays may be used with the immunogen as antigen to assess the levels of antibodies.

Following immunization, antisera reactive with a polypeptide of the invention may be obtained and, if desired, polyclonal antibodies isolated from the serum. To produce monoclonal antibodies, antibody producing cells (lymphocytes) may be harvested from an immunized animal and fused by standard somatic cell fusion procedures with immortalizing cells such as myeloma cells to yield hybridoma cells. Such techniques are well known in the art, and include, for example, the hybridoma technique (originally developed by Kohler and Milstein, (1975) *Nature*, 256: 495-497), as the human B cell hybridoma technique (Kozbar et al., (1983) *Immunology Today*, 4: 72), and the EBV-hybridoma technique to

produce human monoclonal antibodies (Cole et al., (1985) *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc. pp. 77-96). Hybridoma cells can be screened immunochemically for production of antibodies specifically reactive with the polypeptides of the invention and the monoclonal antibodies isolated.

Antibodies directed against the polypeptides of the invention can be used to selectively block the action of the polypeptides of the invention. Antibodies against a polypeptide of the invention may be employed to treat infections, particularly bacterial infections and diseases. For example, the present invention contemplates a method for treating a subject suffering from a *P. aeruginosa* related disease or disorder, comprising administering to an animal having the condition a therapeutically effective amount of a purified antibody that binds specifically to a polypeptide of the invention. In another example, the present invention contemplates a method for inhibiting SEQ ID NO: 2 or SEQ ID NO: 4 dependent growth or infectivity of *P. aeruginosa*, comprising contacting *P. aeruginosa* with a purified antibody that binds specifically to a polypeptide of the invention.

In one embodiment, antibodies reactive with a polypeptide of the invention are used in the immunological screening of cDNA libraries constructed in expression vectors, such as λgt11, λgt18-23, λZAP, and λORF8. Messenger libraries of this type, having coding sequences inserted in the correct reading frame and orientation, can produce fusion proteins. For instance, λgt11 will produce fusion proteins whose amino termini consist of ß-galactosidase amino acid sequences and whose carboxy termini consist of a foreign polypeptide. Antigenic epitopes of a polypeptide of the invention can then be detected with antibodies, as, for example, reacting nitrocellulose filters lifted from phage infected bacterial plates with an antibody specific for a polypeptide of the invention. Phage scored by this assay can then be isolated from the infected plate. Thus, homologs of a polypeptide of the invention can be detected and cloned from other sources.

Antibodies may be employed to isolate or to identify clones expressing the polypeptides to purify the polypeptides by affinity chromatography.

In other embodiments, the polypeptides of the invention may be modified so as to increase their immunogenicity. For example, a polypeptide, such as an antigenically or immunologically equivalent derivative, may be associated, for example by conjugation, with an immunogenic carrier protein for example bovine serum albumin (BSA) or keyhole limpet haemocyanin (KLH). Alternatively a multiple antigenic peptide comprising multiple

copies of the protein or polypeptide, or an antigenically or immunologically equivalent polypeptide thereof may be sufficiently antigenic to improve immunogenicity so as to obviate the use of a carrier.

In other embodiments, the antibodies of the invention, or variants thereof, are modified to make them less immunogenic when administered to a subject. For example, if the subject is human, the antibody may be "humanized"; where the complimentarity determining region(s) of the hybridoma-derived antibody has been transplanted into a human monoclonal antibody, for example as described in Jones, P. et al. (1986), Nature 321, 522-525 or Tempest et al. (1991) Biotechnology 9, 266-273. Also, transgenic mice, or other mammals, may be used to express humanized antibodies. Such humanization may be partial or complete.

The use of a nucleic acid of the invention in genetic immunization may employ a suitable delivery method such as direct injection of plasmid DNA into muscles (Wolff et al., Hum Mol Genet 1992, 1:363, Manthorpe et al., Hum. Gene Ther. 1963:4, 419), delivery of DNA complexed with specific protein carriers (Wu et al., J Biol Chem. 1989: 264,16985), coprecipitation of DNA with calcium phosphate (Benvenisty & Reshef, PNAS USA, 1986:83,9551), encapsulation of DNA in various forms of liposomes (Kaneda et al., Science 1989:243,375), particle bombardment (Tang et al., Nature 1992, 356:152, Eisenbraun et al., DNA Cell Biol 1993, 12:791) and *in vivo* infection using cloned retroviral vectors (Seeger et al., PNAS USA 1984:81,5849).

8. Diagnostic Assays

The invention further provides a method for detecting the presence of *P. aeruginosa* in a biological sample. Detection of *P. aeruginosa* in a subject, particularly a mammal, and especially a human, will provide a diagnostic method for diagnosis of a *P. aeruginosa* related disease or disorder. In general, the method involves contacting the biological sample with a compound or an agent capable of detecting a polypeptide of the invention or a nucleic acid of the invention. The term "biological sample" when used in reference to a diagnostic assay is intended to include tissues, cells and biological fluids isolated from a subject, as well as tissues, cells and fluids present within a subject.

The detection method of the invention may be used to detect the presence of *P. aeruginosa* in a biological sample *in vitro* as well as *in vivo*. For example, *in vitro* techniques for detection of a nucleic acid of the invention include Northern hybridizations

and in situ hybridizations. *In vitro* techniques for detection of polypeptides of the invention include enzyme linked immunosorbent assays (ELISAs), Western blots, immunoprecipitations, immunofluorescence, radioimmunoassays and competitive binding assays. Alternatively, polypeptides of the invention can be detected *in vivo* in a subject by introducing into the subject a labeled antibody specific for a polypeptide of the invention. For example, the antibody can be labeled with a radioactive marker whose presence and location in a subject can be detected by standard imaging techniques. It may be possible to use all of the diagnostic methods disclosed herein for pathogens in addition to *P. aeruginosa*.

Nucleic acids for diagnosis may be obtained from an infected individual's cells and tissues, such as bone, blood, muscle, cartilage, and skin. Nucleic acids, e.g., DNA and RNA, may be used directly for detection or may be amplified, e.g., enzymatically by using PCR or other amplification technique, prior to analysis. Using amplification, characterization of the species and strain of prokaryote present in an individual, may be made by an analysis of the genotype of the prokaryote gene. Deletions and insertions can be detected by a change in size of the amplified product in comparison to the genotype of a reference sequence. Point mutations can be identified by hybridizing a nucleic acid, e.g., amplified DNA, to a nucleic acid of the invention, which nucleic acid may be labeled. Perfectly matched sequences can be distinguished from mismatched duplexes by RNase digestion or by differences in melting temperatures. DNA sequence differences may also be detected by alterations in the electrophoretic mobility of the DNA fragments in gels, with or without denaturing agents, or by direct DNA sequencing. See, e.g. Myers et al., Science, 230: 1242 (1985). Sequence changes at specific locations also may be revealed by nuclease protection assays, such as RNase and S1 protection or a chemical cleavage method. See, e.g., Cotton et al., Proc. Natl. Acad. Sci., USA, 85: 4397-4401 (1985).

Agents for detecting a nucleic acid of the invention, e.g., comprising the sequence set forth in SEQ ID NO: 1 or SEQ ID NO: 3, include labeled or labelable nucleic acid probes capable of hybridizing to a nucleic acid of the invention. The nucleic acid probe can comprise, for example, the full length sequence of a nucleic acid of the invention, or an equivalent thereof, or a portion thereof, such as an oligonucleotide of at least 15, 30, 50, 100, 250 or 500 nucleotides in length and sufficient to specifically hybridize under stringent conditions to SEQ ID NO: 1 or SEQ ID NO: 3, or the complement thereof. Agents for detecting a polypeptide of the invention, e.g., comprising an amino acid sequence of SEQ

ID NO: 2 or SEQ ID NO: 4, include labeled or labelable antibodies capable of binding to a polypeptide of the invention. Antibodies may be polyclonal, or alternatively, monoclonal. An intact antibody, or a fragment thereof (e.g., Fab or $F(ab')_2$) can be used. Labeling the probe or antibody also encompasses direct labeling of the probe or antibody by coupling (e.g., physically linking) a detectable substance to the probe or antibody, as well as indirect labeling of the probe or antibody by reactivity with another reagent that is directly labeled. Examples of indirect labeling include detection of a primary antibody using a fluorescently labeled secondary antibody and end-labeling of a DNA probe with biotin such that it can be detected with fluorescently labeled streptavidin.

In certain embodiments, detection of a nucleic acid of the invention in a biological sample involves the use of a probe/primer in a polymerase chain reaction (PCR) (see, e.g. U.S. Pat. Nos. 4,683,195 and 4,683,202), such as anchor PCR or RACE PCR, or, alternatively, in a ligation chain reaction (LCR) (see, e.g., Landegran et al. (1988) Science 241:1077-1080; and Nakazawa et al. (1994) PNAS 91:360-364), the latter of which can be particularly useful for distinguishing between orthologs of polynucleotides of the invention (see Abravaya et al. (1995) Nucleic Acids Res. 23:675-682). This method can include the steps of collecting a sample of cells from a patient, isolating nucleic acid (e.g., genomic, mRNA or both) from the cells of the sample, contacting the nucleic acid sample with one or more primers which specifically hybridize to a nucleic acid of the invention under conditions such that hybridization and amplification of the polynucleotide (if present) occurs, and detecting the presence or absence of an amplification product, or detecting the size of the amplification product and comparing the length to a control sample.

In one aspect, the present invention contemplates a method for detecting the presence of *P. aeruginosa* in a sample, the method comprising: (a) providing a sample to be tested for the presence of *P. aeruginosa*; (b) contacting the sample with an antibody reactive against eight consecutive amino acid residues of SEQ ID NO: 2 or SEQ ID NO: 4 under conditions which permit association between the antibody and its ligand; and (c) detecting interaction of the antibody with its ligand, thereby detecting the presence of *P. aeruginosa* in the sample.

In another aspect, the present invention contemplates a method for detecting the presence of *P. aeruginosa* in a sample, the method comprising: (a) providing a sample to be tested for the presence of *P. aeruginosa*; (b) contacting the sample with an antibody that binds specifically to a polypeptide of the invention under conditions which permit

association between the antibody and its ligand; and (c) detecting interaction of the antibody with its ligand, thereby detecting the presence of *P. aeruginosa* in the sample.

In yet another example, the present invention contemplates a method for diagnosing a patient suffering from a *P. aeruginosa* related disease or disorder, comprising: (a) obtaining a biological sample from a patient; (b) detecting the presence or absence of a polypeptide of the invention, or a nucleic acid encoding a polypeptide of the invention, in the sample; and (c) diagnosing a patient suffering from a *P. aeruginosa* related disease or disorder based on the presence of a polypeptide of the invention, or a nucleic acid encoding a polypeptide of the invention, in the patient sample.

The diagnostic assays of the invention may also be used to monitor the effectiveness of an anti-*P. aeruginosa* treatment in an individual suffering from an *P. aeruginosa* related disease or disorder. For example, the presence and/or amount of a nucleic acid of the invention or a polypeptide of the invention can be detected in an individual suffering from an *P. aeruginosa* related disease or disorder before and after treatment with anti-*P. aeruginosa* therapeutic agent. Any change in the level of a polynucleotide or polypeptide of the invention after treatment of the individual with the therapeutic agent can provide information about the effectiveness of the treatment course. In particular, no change, or a decrease, in the level of a polynucleotide or polypeptide of the invention present in the biological sample will indicate that the therapeutic is successfully combating the *P. aeruginosa* related disease or disorder.

The invention also encompasses kits for detecting the presence of *P. aeruginosa* in a biological sample. For example, the kit can comprise a labeled or labelable compound or agent capable of detecting a polynucleotide or polypeptide of the invention in a biological sample; means for determining the amount of *P. aeruginosa* in the sample; and means for comparing the amount of *P. aeruginosa* in the sample with a standard. The compound or agent can be packaged in a suitable container. The kit can further comprise instructions for using the kit to detect a polynucleotide or polypeptide of the invention.

9. Drug Discovery

Modulators to polypeptides of the invention and other structurally related molecules, and complexes containing the same, may be identified and developed as set forth below and otherwise using techniques and methods known to those of skill in the art. The modulators of the invention may be employed, for instance, to inhibit and treat *P.*

aeruginosa associated diseases or conditions, such as osteomyelitis, otitis externa, conjunctivitis, keratitis, endophthalmitis, alveolar necrosis, vascular invasion, bacteremia, and burn infection.

A variety of methods for inhibiting the growth or infectivity of *P. aeruginosa* are contemplated by the present invention. For example, exemplary methods involve contacting *P. aeruginosa* with a polypeptide of the invention that modulates the same or another polypeptide from such pathogen, a nucleic acid encoding such polypeptide of the invention, or a compound thought or shown to be effective against such pathogen.

For example, in one aspect, the present invention contemplates a method for treating a patient suffering from an infection of *P. aeruginosa*, comprising administering to the patient an amount of a SEQ ID NO: 2 or SEQ ID NO: 4 inhibitor effective to inhibit the expression and/or activity of a polypeptide of the invention. In certain instances, the animal is a human or a livestock animal such as a cow, pig, goat or sheep. The present invention further contemplates a method for treating a subject suffering from a *P. aeruginosa* related disease or disorder, comprising administering to an animal having the condition a therapeutically effective amount of a molecule identified using one of the methods of the present invention.

The present invention contemplates making any molecule that is shown to modulate the activity of a polypeptide of the invention.

In another embodiment, inhibitors, modulators of the subject polypeptides, or biological complexes containing them, may be used in the manufacture of a medicament for any number of uses, including, for example, treating any disease or other treatable condition of a patient (including humans and animals), and particularly a disease caused by *P. aeruginosa*, such as, for example, one of the following: osteomyelitis, otitis externa, conjunctivitis, keratitis, endophthalmitis, alveolar necrosis, vascular invasion, bacteremia, and burn infection.

(a) Drug Design

A number of techniques can be used to screen, identify, select and design chemical entities capable of associating with polypeptides of the invention, structurally homologous molecules, and other molecules. Knowledge of the structure for a polypeptide of the invention, determined in accordance with the methods described herein, permits the design and/or identification of molecules and/or other modulators which have a shape complementary to the conformation of a polypeptide of the invention, or more particularly,

a druggable region thereof. It is understood that such techniques and methods may use, in addition to the exact structural coordinates and other information for a polypeptide of the invention, structural equivalents thereof described above (including, for example, those structural coordinates that are derived from the structural coordinates of amino acids contained in a druggable region as described above).

The term "chemical entity," as used herein, refers to chemical compounds, complexes of two or more chemical compounds, and fragments of such compounds or complexes. In certain instances, it is desirable to use chemical entities exhibiting a wide range of structural and functional diversity, such as compounds exhibiting different shapes (e.g., flat aromatic rings(s), puckered aliphatic rings(s), straight and branched chain aliphatics with single, double, or triple bonds) and diverse functional groups (e.g., carboxylic acids, esters, ethers, amines, aldehydes, ketones, and various heterocyclic rings).

In one aspect, the method of drug design generally includes computationally evaluating the potential of a selected chemical entity to associate with any of the molecules or complexes of the present invention (or portions thereof). For example, this method may include the steps of (a) employing computational means to perform a fitting operation between the selected chemical entity and a druggable region of the molecule or complex; and (b) analyzing the results of said fitting operation to quantify the association between the chemical entity and the druggable region.

A chemical entity may be examined either through visual inspection or through the use of computer modeling using a docking program such as GRAM, DOCK, or AUTODOCK (Dunbrack et al., Folding & Design, 2:27-42 (1997)). This procedure can include computer fitting of chemical entities to a target to ascertain how well the shape and the chemical structure of each chemical entity will complement or interfere with the structure of the subject polypeptide (Bugg et al., Scientific American, Dec.: 92-98 (1993); West et al., TIPS, 16:67-74 (1995)). Computer programs may also be employed to estimate the attraction, repulsion, and steric hindrance of the chemical entity to a druggable region, for example. Generally, the tighter the fit (e.g., the lower the steric hindrance, and/or the greater the attractive force) the more potent the chemical entity will be because these properties are consistent with a tighter binding constant. Furthermore, the more specificity in the design of a chemical entity the more likely that the chemical entity will not interfere with related proteins, which may minimize potential side-effects due to unwanted interactions.

A variety of computational methods for molecular design, in which the steric and electronic properties of druggable regions are used to guide the design of chemical entities, are known: Cohen et al. (1990) *J. Med. Cam.* 33: 883-894; Kuntz et al. (1982) *J. Mol. Biol* 161: 269-288; DesJarlais (1988) *J. Med. Cam.* 31: 722-729; Bartlett et al. (1989) *Spec. Publ., Roy. Soc. Chem.* 78: 182-196; Goodford et al. (1985) *J. Med. Cam.* 28: 849-857; and DesJarlais et al. *J. Med. Cam.* 29: 2149-2153. Directed methods generally fall into two categories: (1) design by analogy in which 3-D structures of known chemical entities (such as from a crystallographic database) are docked to the druggable region and scored for goodness-of-fit; and (2) *de novo* design, in which the chemical entity is constructed piece-wise in the druggable region. The chemical entity may be screened as part of a library or a database of molecules. Databases which may be used include ACD (Molecular Designs Limited), NCI (National Cancer Institute), CCDC (Cambridge Crystallographic Data Center), CAST (Chemical Abstract Service), Derwent (Derwent Information Limited), Maybridge (Maybridge Chemical Company Ltd), Aldrich (Aldrich Chemical Company), DOCK (University of California in San Francisco), and the Directory of Natural Products (Chapman & Hall). Computer programs such as CONCORD (Tripos Associates) or DB-Converter (Molecular Simulations Limited) can be used to convert a data set represented in two dimensions to one represented in three dimensions.

Chemical entities may be tested for their capacity to fit spatially with a druggable region or other portion of a target protein. As used herein, the term "fits spatially" means that the three-dimensional structure of the chemical entity is accommodated geometrically by a druggable region. A favorable geometric fit occurs when the surface area of the chemical entity is in close proximity with the surface area of the druggable region without forming unfavorable interactions. A favorable complementary interaction occurs where the chemical entity interacts by hydrophobic, aromatic, ionic, dipolar, or hydrogen donating and accepting forces. Unfavorable interactions may be steric hindrance between atoms in the chemical entity and atoms in the druggable region.

If a model of the present invention is a computer model, the chemical entities may be positioned in a druggable region through computational docking. If, on the other hand, the model of the present invention is a structural model, the chemical entities may be positioned in the druggable region by, for example, manual docking. As used herein the term "docking" refers to a process of placing a chemical entity in close proximity with a

druggable region, or a process of finding low energy conformations of a chemical entity/druggable region complex.

In an illustrative embodiment, the design of potential modulator begins from the general perspective of shape complimentary for the druggable region of a polypeptide of the invention, and a search algorithm is employed which is capable of scanning a database of small molecules of known three-dimensional structure for chemical entities which fit geometrically with the target druggable region. Most algorithms of this type provide a method for finding a wide assortment of chemical entities that are complementary to the shape of a druggable region of the subject polypeptide. Each of a set of chemical entities from a particular data-base, such as the Cambridge Crystallographic Data Bank (CCDB) (Allen et al. (1973) *J. Chem. Doc.* 13: 119), is individually docked to the druggable region of a polypeptide of the invention in a number of geometrically permissible orientations with use of a docking algorithm. In certain embodiments, a set of computer algorithms called DOCK, can be used to characterize the shape of invaginations and grooves that form the active sites and recognition surfaces of the druggable region (Kuntz et al. (1982) *J. Mol. Biol* 161: 269-288). The program can also search a database of small molecules for templates whose shapes are complementary to particular binding sites of a polypeptide of the invention (DesJarlais et al. (1988) *J Med Chem* 31: 722-729).

The orientations are evaluated for goodness-of-fit and the best are kept for further examination using molecular mechanics programs, such as AMBER or CHARMM. Such algorithms have previously proven successful in finding a variety of chemical entities that are complementary in shape to a druggable region.

Goodford (1985, *J Med Chem* 28:849-857) and Boobbyer et al. (1989, *J Med Chem* 32:1083-1094) have produced a computer program (GRID) which seeks to determine regions of high affinity for different chemical groups (termed probes) of the druggable region. GRID hence provides a tool for suggesting modifications to known chemical entities that might enhance binding. It may be anticipated that some of the sites discerned by GRID as regions of high affinity correspond to "pharmacophoric patterns" determined inferentially from a series of known ligands. As used herein, a "pharmacophoric pattern" is a geometric arrangement of features of chemical entities that is believed to be important for binding. Attempts have been made to use pharmacophoric patterns as a search screen for novel ligands (Jakes et al. (1987) *J Mol Graph* 5:41-48; Brint et al. (1987) *J Mol Graph* 5:49-56; Jakes et al. (1986) *J Mol Graph* 4:12-20).

Yet a further embodiment of the present invention utilizes a computer algorithm such as CLIX which searches such databases as CCDB for chemical entities which can be oriented with the druggable region in a way that is both sterically acceptable and has a high likelihood of achieving favorable chemical interactions between the chemical entity and the surrounding amino acid residues. The method is based on characterizing the region in terms of an ensemble of favorable binding positions for different chemical groups and then searching for orientations of the chemical entities that cause maximum spatial coincidence of individual candidate chemical groups with members of the ensemble. The algorithmic details of CLIX is described in Lawrence et al. (1992) *Proteins* 12:31-41.

In this way, the efficiency with which a chemical entity may bind to or interfere with a druggable region may be tested and optimized by computational evaluation. For example, for a favorable association with a druggable region, a chemical entity must preferably demonstrate a relatively small difference in energy between its bound and fine states (i.e., a small deformation energy of binding). Thus, certain, more desirable chemical entities will be designed with a deformation energy of binding of not greater than about 10 kcal/mole, and more preferably, not greater than 7 kcal/mole. Chemical entities may interact with a druggable region in more than one conformation that is similar in overall binding energy. In those cases, the deformation energy of binding is taken to be the difference between the energy of the free entity and the average energy of the conformations observed when the chemical entity binds to the target.

In this way, the present invention provides computer-assisted methods for identifying or designing a potential modulator of the activity of a polypeptide of the invention including: supplying a computer modeling application with a set of structure coordinates of a molecule or complex, the molecule or complex including at least a portion of a druggable region from a polypeptide of the invention; supplying the computer modeling application with a set of structure coordinates of a chemical entity; and determining whether the chemical entity is expected to bind to the molecule or complex, wherein binding to the molecule or complex is indicative of potential modulation of the activity of a polypeptide of the invention.

In another aspect, the present invention provides a computer-assisted method for identifying or designing a potential modulator to a polypeptide of the invention, supplying a computer modeling application with a set of structure coordinates of a molecule or complex, the molecule or complex including at least a portion of a druggable region of a

polypeptide of the invention; supplying the computer modeling application with a set of structure coordinates for a chemical entity; evaluating the potential binding interactions between the chemical entity and active site of the molecule or molecular complex; structurally modifying the chemical entity to yield a set of structure coordinates for a modified chemical entity, and determining whether the modified chemical entity is expected to bind to the molecule or complex, wherein binding to the molecule or complex is indicative of potential modulation of the polypeptide of the invention.

In one embodiment, a potential modulator can be obtained by screening a peptide library (Scott and Smith, Science, 249:386-390 (1990); Cwirla et al., Proc. Natl. Acad. Sci., 87:6378-6382 (1990); Devlin et al., Science, 249:404-406 (1990)). A potential modulator selected in this manner could then be systematically modified by computer modeling programs until one or more promising potential drugs are identified. Such analysis has been shown to be effective in the development of HIV protease inhibitors (Lam et al., Science 263:380-384 (1994); Wlodawer et al., Ann. Rev. Biochem. 62:543-585 (1993); Appelt, Perspectives in Drug Discovery and Design 1:23-48 (1993); Erickson, Perspectives in Drug Discovery and Design 1:109-128 (1993)). Alternatively a potential modulator may be selected from a library of chemicals such as those that can be licensed from third parties, such as chemical and pharmaceutical companies. A third alternative is to synthesize the potential modulator *de novo*.

For example, in certain embodiments, the present invention provides a method for making a potential modulator for a polypeptide of the invention, the method including synthesizing a chemical entity or a molecule containing the chemical entity to yield a potential modulator of a polypeptide of the invention, the chemical entity having been identified during a computer-assisted process including supplying a computer modeling application with a set of structure coordinates of a molecule or complex, the molecule or complex including at least one druggable region from a polypeptide of the invention; supplying the computer modeling application with a set of structure coordinates of a chemical entity; and determining whether the chemical entity is expected to bind to the molecule or complex at the active site, wherein binding to the molecule or complex is indicative of potential modulation. This method may further include the steps of evaluating the potential binding interactions between the chemical entity and the active site of the molecule or molecular complex and structurally modifying the chemical entity to yield a set

of structure coordinates for a modified chemical entity, which steps may be repeated one or more times.

Once a potential modulator is identified, it can then be tested in any standard assay for the macromolecule depending of course on the macromolecule, including in high throughput assays. Further refinements to the structure of the modulator will generally be necessary and can be made by the successive iterations of any and/or all of the steps provided by the particular screening assay, in particular further structural analysis by e.g., ^{15}N NMR relaxation rate determinations or x-ray crystallography with the modulator bound to the subject polypeptide. These studies may be performed in conjunction with biochemical assays.

Once identified, a potential modulator may be used as a model structure, and analogs to the compound can be obtained. The analogs are then screened for their ability to bind the subject polypeptide. An analog of the potential modulator might be chosen as a modulator when it binds to the subject polypeptide with a higher binding affinity than the predecessor modulator.

In a related approach, iterative drug design is used to identify modulators of a target protein. Iterative drug design is a method for optimizing associations between a protein and a modulator by determining and evaluating the three dimensional structures of successive sets of protein/modulator complexes. In iterative drug design, crystals of a series of protein/modulator complexes are obtained and then the three-dimensional structures of each complex is solved. Such an approach provides insight into the association between the proteins and modulators of each complex. For example, this approach may be accomplished by selecting modulators with inhibitory activity, obtaining crystals of this new protein/modulator complex, solving the three dimensional structure of the complex, and comparing the associations between the new protein/modulator complex and previously solved protein/modulator complexes. By observing how changes in the modulator affected the protein/modulator associations, these associations may be optimized.

In addition to designing and/or identifying a chemical entity to associate with a druggable region, as described above, the same techniques and methods may be used to design and/or identify chemical entities that either associate, or do not associate, with affinity regions, selectivity regions or undesired regions of protein targets. By such methods, selectivity for one or a few targets, or alternatively for multiple targets, from the same species or from multiple species, can be achieved.

For example, a chemical entity may be designed and/or identified for which the binding energy for one druggable region, e.g., an affinity region or selectivity region, is more favorable than that for another region, e.g., an undesired region, by about 20%, 30%, 50% to about 60% or more. It may be the case that the difference is observed between (a) more than two regions, (b) between different regions (selectivity, affinity or undesirable) from the same target, (c) between regions of different targets, (d) between regions of homologs from different species, or (e) between other combinations. Alternatively, the comparison may be made by reference to the Kd, usually the apparent Kd, of said chemical entity with the two or more regions in question.

In another aspect, prospective modulators are screened for binding to two nearby druggable regions on a target protein. For example, a modulator that binds a first region of a target polypeptide does not bind a second nearby region. Binding to the second region can be determined by monitoring changes in a different set of amide chemical shifts in either the original screen or a second screen conducted in the presence of a modulator (or potential modulator) for the first region. From an analysis of the chemical shift changes, the approximate location of a potential modulator for the second region is identified. Optimization of the second modulator for binding to the region is then carried out by screening structurally related compounds (e.g., analogs as described above). When modulators for the first region and the second region are identified, their location and orientation in the ternary complex can be determined experimentally. On the basis of this structural information, a linked compound, e.g., a consolidated modulator, is synthesized in which the modulator for the first region and the modulator for the second region are linked. In certain embodiments, the two modulators are covalently linked to form a consolidated modulator. This consolidated modulator may be tested to determine if it has a higher binding affinity for the target than either of the two individual modulators. A consolidated modulator is selected as a modulator when it has a higher binding affinity for the target than either of the two modulators. Larger consolidated modulators can be constructed in an analogous manner, e.g., linking three modulators which bind to three nearby regions on the target to form a multilinked consolidated modulator that has an even higher affinity for the target than the linked modulator. In this example, it is assumed that is desirable to have the modulator bind to all the druggable regions. However, it may be the case that binding to certain of the druggable regions is not desirable, so that the same techniques may be used to

identify modulators and consolidated modulators that show increased specificity based on binding to at least one but not all druggable regions of a target.

The present invention provides a number of methods that use drug design as described above. For example, in one aspect, the present invention contemplates a method for designing a candidate compound for screening for inhibitors of a polypeptide of the invention, the method comprising: (a) determining the three dimensional structure of a crystallized polypeptide of the invention or a fragment thereof; and (b) designing a candidate inhibitor based on the three dimensional structure of the crystallized polypeptide or fragment.

In another aspect, the present invention contemplates a method for identifying a potential inhibitor of a polypeptide of the invention, the method comprising: (a) providing the three-dimensional coordinates of a polypeptide of the invention or a fragment thereof; (b) identifying a druggable region of the polypeptide or fragment; and (c) selecting from a database at least one compound that comprises three dimensional coordinates which indicate that the compound may bind the druggable region; (d) wherein the selected compound is a potential inhibitor of a polypeptide of the invention.

In another aspect, the present invention contemplates a method for identifying a potential modulator of a molecule comprising a druggable region similar to that of SEQ ID NO: 2 or SEQ ID NO: 4, the method comprising: (a) using the atomic coordinates of amino acid residues from SEQ ID NO: 2 or SEQ ID NO: 4, or a fragment thereof, ± a root mean square deviation from the backbone atoms of the amino acids of not more than 1.5 Å, to generate a three-dimensional structure of a molecule comprising a druggable region that is a portion of SEQ ID NO: 2 or SEQ ID NO: 4; (b) employing the three dimensional structure to design or select the potential modulator; (c) synthesizing the modulator; and (d) contacting the modulator with the molecule to determine the ability of the modulator to interact with the molecule.

In another aspect, the present invention contemplates an apparatus for determining whether a compound is a potential inhibitor of a polypeptide having SEQ ID NO: 2 or SEQ ID NO: 4, the apparatus comprising: (a) a memory that comprises: (i) the three dimensional coordinates and identities of the atoms of a polypeptide of the invention or a fragment thereof that form a druggable site; and (ii) executable instructions; and (b) a processor that is capable of executing instructions to: (i) receive three-dimensional structural information for a candidate compound; (ii) determine if the three-dimensional structure of the candidate

compound is complementary to the structure of the interior of the druggable site; and (iii) output the results of the determination.

In another aspect, the present invention contemplates a method for designing a potential compound for the prevention or treatment of *P. aeruginosa* related disease or disorder, the method comprising: (a) providing the three dimensional structure of a crystallized polypeptide of the invention, or a fragment thereof; (b) synthesizing a potential compound for the prevention or treatment of *P. aeruginosa* related disease or disorder based on the three dimensional structure of the crystallized polypeptide or fragment; (c) contacting a polypeptide of the present invention or an *P. aeruginosa* with the potential compound; and (d) assaying the activity of a polypeptide of the present invention, wherein a change in the activity of the polypeptide indicates that the compound may be useful for prevention or treatment of a *P. aeruginosa* related disease or disorder.

In another aspect, the present invention contemplates a method for designing a potential compound for the prevention or treatment of *P. aeruginosa* related disease or disorder, the method comprising: (a) providing structural information of a druggable region derived from NMR spectroscopy of a polypeptide of the invention, or a fragment thereof; (b) synthesizing a potential compound for the prevention or treatment of *P. aeruginosa* related disease or disorder based on the structural information; (c) contacting a polypeptide of the present invention or an *P. aeruginosa* with the potential compound; and (d) assaying the activity of a polypeptide of the present invention, wherein a change in the activity of the polypeptide indicates that the compound may be useful for prevention or treatment of a *P. aeruginosa* related disease or disorder.

(b) In Vitro Assays

Polypeptides of the invention may be used to assess the activity of small molecules and other modulators in *in vitro* assays. In one embodiment of such an assay, agents are identified which modulate the biological activity of a protein, protein-protein interaction of interest or protein complex, such as an enzymatic activity, binding to other cellular components, cellular compartmentalization, signal transduction, and the like. In certain embodiments, the test agent is a small organic molecule.

Assays may employ kinetic or thermodynamic methodology using a wide variety of techniques including, but not limited to, microcalorimetry, circular dichroism, capillary zone electrophoresis, nuclear magnetic resonance spectroscopy, fluorescence spectroscopy, and combinations thereof.

The invention also provides a method of screening compounds to identify those which modulate the action of polypeptides of the invention, or polynucleotides encoding the same. The method of screening may involve high-throughput techniques. For example, to screen for modulators, a synthetic reaction mix, a cellular compartment, such as a membrane, cell envelope or cell wall, or a preparation of any thereof, comprising a polypeptide of the invention and a labeled substrate or ligand of such polypeptide is incubated in the absence or the presence of a candidate molecule that may be a modulator of a polypeptide of the invention. The ability of the candidate molecule to modulate a polypeptide of the invention is reflected in decreased binding of the labeled ligand or decreased production of product from such substrate. Detection of the rate or level of production of product from substrate may be enhanced by using a reporter system. Reporter systems that may be useful in this regard include but are not limited to colorimetric labeled substrate converted into product, a reporter gene that is responsive to changes in a nucleic acid of the invention or polypeptide activity, and binding assays known in the art.

Another example of an assay for a modulator of a polypeptide of the invention is a competitive assay that combines a polypeptide of the invention and a potential modulator with molecules that bind to a polypeptide of the invention, recombinant molecules that bind to a polypeptide of the invention, natural substrates or ligands, or substrate or ligand mimetics, under appropriate conditions for a competitive inhibition assay. Polypeptides of the invention can be labeled, such as by radioactivity or a colorimetric compound, such that the number of molecules of a polypeptide of the invention bound to a binding molecule or converted to product can be determined accurately to assess the effectiveness of the potential modulator.

A number of methods for identifying a molecule which modulates the activity of a polypeptide are known in the art. For example, in one such method, a subject polypeptide is contacted with a test compound, and the activity of the subject polypeptide in the presence of the test compound is determined, wherein a change in the activity of the subject polypeptide is indicative that the test compound modulates the activity of the subject polypeptide. In certain instances, the test compound agonizes the activity of the subject polypeptide, and in other instances, the test compound antagonizes the activity of the subject polypeptide.

In another example, a compound which modulates SEQ ID NO: 2 or SEQ ID NO: 4 dependent growth or infectivity of *P. aeruginosa* may be identified by (a) contacting a polypeptide of the invention with a test compound; and (b) determining the activity of the polypeptide in the presence of the test compound, wherein a change in the activity of the polypeptide is indicative that the test compound may modulate the growth or infectivity of *P. aeruginosa*.

(c) In Vivo Assays

Animal models of bacterial infection and/or disease may be used as an *in vivo* assay for evaluating the effectiveness of a potential drug target in treating or preventing diseases or disorders. A number of suitable animal models are described briefly below, however, these models are only examples and modifications, or completely different animal models, may be used in accord with the methods of the invention.

(i) Mouse Soft Tissue Model

The mouse soft tissue infection model is a sensitive and effective method for measurement of bacterial proliferation. In these models (Vogelman et al., 1988, J. Infect. Dis. 157: 287-298) anesthetized mice are infected with the bacteria in the muscle of the hind thigh. The mice can be either chemically immune compromised (e.g., cytoxan treated at 125 mg/kg on days -4, -2, and 0) or immunocompetent. The dose of microbe necessary to cause an infection is variable and depends on the individual microbe, but commonly is on the order of 10^5 - 10^6 colony forming units per injection for bacteria. A variety of mouse strains are useful in this model although Swiss Webster and DBA2 lines are most commonly used. Once infected the animals are conscious and show no overt ill effects of the infections for approximately 12 hours. After that time virulent strains cause swelling of the thigh muscle, and the animals can become bacteremic within approximately 24 hours. This model most effectively measures proliferation of the microbe, and this proliferation is measured by sacrifice of the infected animal and counting colonies from homogenized thighs.

(ii) Diffusion Chamber Model

A second model useful for assessing the virulence of microbes is the diffusion chamber model (Malouin et al., 1990, Infect. Immun. 58: 1247-1253; Doy et al., 1980, J. Infect. Dis. 2: 39-51; Kelly et al., 1989, Infect. Immun. 57: 344-350. In this model rodents have a diffusion chamber surgically placed in the peritoneal cavity. The chamber consists of a polypropylene cylinder with semipermeable membranes covering the chamber ends.

Diffusion of peritoneal fluid into and out of the chamber provides nutrients for the microbes. The progression of the "infection" may be followed by examining growth, the exoproduct production or RNA messages. The time experiments are done by sampling multiple chambers.

(iii) Endocarditis Model

For bacteria, an important animal model effective in assessing pathogenicity and virulence is the endocarditis model (J. Santoro and M. E. Levinson, 1978, Infect. Immun. 19: 915-918). A rat endocarditis model can be used to assess colonization, virulence and proliferation.

(iv) Osteomyelitis Model

A fourth model useful in the evaluation of pathogenesis is the osteomyelitis model (Spagnolo et al., 1993, Infect. Immun. 61: 5225-5230). Rabbits are used for these experiments. Anesthetized animals have a small segment of the tibia removed and microorganisms are microinjected into the wound. The excised bone segment is replaced and the progression of the disease is monitored. Clinical signs, particularly inflammation and swelling are monitored. Termination of the experiment allows histolic and pathologic examination of the infection site to complement the assessment procedure.

(v) Murine Septic Arthritis Model

A fifth model relevant to the study of microbial pathogenesis is a murine septic arthritis model (Abdelnour et al., 1993, Infect. Immun. 61: 3879-3885). In this model mice are infected intravenously and pathogenic organisms are found to cause inflammation in distal limb joints. Monitoring of the inflammation and comparison of inflammation vs. inocula allows assessment of the virulence of related strains.

(vi) Bacterial Peritonitis Model

Finally, bacterial peritonitis offers rapid and predictive data on the virulence of strains (M. G. Bergeron, 1978, Scand. J. Infect. Dis. Suppl. 14: 189-206; S. D. Davis, 1975, Antimicrob. Agents Chemother. 8: 50-53). Peritonitis in rodents, such as mice, can provide essential data on the importance of targets. The end point may be lethality or clinical signs can be monitored. Variation in infection dose in comparison to outcome allows evaluation of the virulence of individual strains.

A variety of other *in vivo* models are available and may be used when appropriate for specific pathogens or specific test agents. For example, target organ recovery assays (Gordee et al., 1984, J. Antibiotics 37:1054-1065; Bannatyne et al., 1992, Infect. 20:168-

170) may be useful for fungi and for bacterial pathogens which are not acutely virulent to animals.

It is also relevant to note that the species of animal used for an infection model, and the specific genetic make-up of that animal, may contribute to the effective evaluation of the effects of a particular test agent. For example, immuno-incompetent animals may, in some instances, be preferable to immuno-competent animals. For example, the action of a competent immune system may, to some degree, mask the effects of the test agent as compared to a similar infection in an immuno-incompetent animal. In addition, many opportunistic infections, in fact, occur in immuno-compromised patients, so modeling an infection in a similar immunological environment is appropriate.

10. Vaccines

There are provided by the invention, products, compositions and methods for raising immunological response against a pathogen, especially *P. aeruginosa*. In one aspect, a polypeptide of the invention or a nucleic acid of the invention, or an antigenic fragment thereof, may be administered to a subject, optionally with a booster, adjuvant, or other composition that stimulates immune responses.

Another aspect of the invention relates to a method for inducing an immunological response in an individual, particularly a mammal which comprises inoculating the individual with a polypeptide of the invention and/or a nucleic acid of the invention, adequate to produce antibody and/or T cell immune response to protect said individual from infection, particularly bacterial infection and most particularly *P. aeruginosa* infection. Also provided are methods whereby such immunological response slows bacterial replication. Yet another aspect of the invention relates to a method of inducing immunological response in an individual which comprises delivering to such individual a nucleic acid vector, sequence or ribozyme to direct expression of a polypeptide of the invention and/or a nucleic acid of the invention *in vivo* in order to induce an immunological response, such as, to produce antibody and/or T cell immune response, including, for example, cytokine-producing T cells or cytotoxic T cells, to protect said individual, preferably a human, from disease, whether that disease is already established within the individual or not. One example of administering the gene is by accelerating it into the desired cells as a coating on particles or otherwise. Such nucleic acid vector may comprise

DNA, RNA, a ribozyme, a modified nucleic acid, a DNA/RNA hybrid, a DNA-protein complex or an RNA-protein complex.

A further aspect of the invention relates to an immunological composition that when introduced into an individual, preferably a human, capable of having induced within it an immunological response, induces an immunological response in such individual to a nucleic acid of the invention and/or a polypeptide encoded therefrom, wherein the composition comprises a recombinant nucleic acid of the invention and/or polypeptide encoded therefrom and/or comprises DNA and/or RNA which encodes and expresses an antigen of said nucleic acid of the invention, polypeptide encoded therefrom, or other polypeptide of the invention. The immunological response may be used therapeutically or prophylactically and may take the form of antibody immunity and/or cellular immunity, such as cellular immunity arising from CTL or CD4+T cells.

In another embodiment, the invention relates to compositions comprising a polypeptide of the invention and an adjuvant. The adjuvant can be any vehicle which would typically enhance the antigenicity of a polypeptide, e.g., minerals (for instance, alum, aluminum hydroxide or aluminum phosphate), saponins complexed to membrane protein antigens (immune stimulating complexes), pluronic polymers with mineral oil, killed mycobacteria in mineral oil, Freund's complete adjuvant, bacterial products, such as muramyl dipeptide (MDP) and lipopolysaccharide (LPS), as well as lipid A, liposomes, or any of the other adjuvants known in the art. A polypeptide of the invention can be emulsified with, absorbed onto, or coupled with the adjuvant.

A polypeptide of the invention may be fused with co-protein or chemical moiety which may or may not by itself produce antibodies, but which is capable of stabilizing the first protein and producing a fused or modified protein which will have antigenic and/or immunogenic properties, and preferably protective properties. Thus fused recombinant protein, may further comprise an antigenic co-protein, such as lipoprotein D from Hemophilus influenzae, Glutathione-S-transferase (GST) or beta-galactosidase, or any other relatively large co-protein which solubilizes the protein and facilitates production and purification thereof. Moreover, the co-protein may act as an adjuvant in the sense of providing a generalized stimulation of the immune system of the organism receiving the protein. The co-protein may be attached to either the amino- or carboxy-terminus of a polypeptide of the invention.

Provided by this invention are compositions, particularly vaccine compositions, and methods comprising the polypeptides and/or polynucleotides of the invention and immunostimulatory DNA sequences, such as those described in Sato, Y. et al. Science 273: 352 (1996).

Also, provided by this invention are methods using the described polynucleotide or particular fragments thereof, which have been shown to encode non-variable regions of bacterial cell surface proteins, in polynucleotide constructs used in such genetic immunization experiments in animal models of infection with *P. aeruginosa*. Such experiments will be particularly useful for identifying protein epitopes able to provoke a prophylactic or therapeutic immune response. It is believed that this approach will allow for the subsequent preparation of monoclonal antibodies of particular value, derived from the requisite organ of the animal successfully resisting or clearing infection, for the development of prophylactic agents or therapeutic treatments of bacterial infection, particularly *P. aeruginosa* infection, in mammals, particularly humans.

A polypeptide of the invention may be used as an antigen for vaccination of a host to produce specific antibodies which protect against invasion of bacteria, for example by blocking adherence of bacteria to damaged tissue.

11. Array Analysis

In part, the present invention is directed to the use of subject nucleic acids in arrays to assess gene expression. In another part, the present invention is directed to the use of subject nucleic acids in arrays for *P. aeruginosa*. In yet another part, the present invention contemplates using the subject nucleic acids to interact with probes contained on arrays.

In one aspect, the present invention contemplates an array comprising a substrate having a plurality of addresses, wherein at least one of the addresses has disposed thereon a capture probe that can specifically bind to a nucleic acid of the invention. In another aspect, the present invention contemplates a method for detecting expression of a nucleotide sequence which encodes a polypeptide of the invention, or a fragment thereof, using the foregoing array by: (a) providing a sample comprising at least one mRNA molecule; (b) exposing the sample to the array under conditions which promote hybridization between the capture probe disposed on the array and a nucleic acid complementary thereto; and (c) detecting hybridization between an mRNA molecule of the

sample and the capture probe disposed on the array, thereby detecting expression of a sequence which encodes for a polypeptide of the invention, or a fragment thereof.

Arrays are often divided into microarrays and macroarrays, where microarrays have a much higher density of individual probe species per area. Microarrays may have as many as 1000 or more different probes in a 1 cm^2 area. There is no concrete cut-off to demarcate the difference between micro- and macroarrays, and both types of arrays are contemplated for use with the invention.

Microarrays are known in the art and generally consist of a surface to which probes that correspond in sequence to gene products (e.g., cDNAs, mRNAs, oligonucleotides) are bound at known positions. In one embodiment, the microarray is an array (e.g., a matrix) in which each position represents a discrete binding site for a product encoded by a gene (e.g., a protein or RNA), and in which binding sites are present for products of most or almost all of the genes in the organism's genome. In certain embodiments, the binding site or site is a nucleic acid or nucleic acid analogue to which a particular cognate cDNA can specifically hybridize. The nucleic acid or analogue of the binding site may be, e.g., a synthetic oligomer, a full-length cDNA, a less-than full length cDNA, or a gene fragment.

Although in certain embodiments the microarray contains binding sites for products of all or almost all genes in the target organism's genome, such comprehensiveness is not necessarily required. Usually the microarray will have binding sites corresponding to at least 100, 500, 1000, 4000 genes or more. In certain embodiments, arrays will have anywhere from about 50, 60, 70 , 80, 90, or even more than 95% of the genes of a particular organism represented. The microarray typically has binding sites for genes relevant to testing and confirming a biological network model of interest. Several exemplary human microarrays are publicly available.

The probes to be affixed to the arrays are typically polynucleotides. These DNAs can be obtained by, e.g., polymerase chain reaction (PCR) amplification of gene segments from genomic DNA, cDNA (e.g., by RT-PCR), or cloned sequences. PCR primers are chosen, based on the known sequence of the genes or cDNA, that result in amplification of unique fragments (e.g., fragments that do not share more than 10 bases of contiguous identical sequence with any other fragment on the microarray). Computer programs are useful in the design of primers with the required specificity and optimal amplification properties. See, e.g., Oligo pl version 5.0 (National Biosciences). In an alternative embodiment, the binding (hybridization) sites are made from plasmid or phage clones of

genes, cDNAs (e.g., expressed sequence tags), or inserts therefrom (Nguyen et al., 1995, Genomics 29:207-209).

A number of methods are known in the art for affixing the nucleic acids or analogues to a solid support that makes up the array (Schena et al., 1995, Science 270:467-470; DeRisi et al., 1996, Nature Genetics 14:457-460; Shalon et al., 1996, Genome Res. 6:639-645; and Schena et al., 1995, Proc. Natl. Acad. Sci. USA 93:10539-11286).

Another method for making microarrays is by making high-density oligonucleotide arrays (Fodor et al., 1991, Science 251:767-773; Pease et al., 1994, Proc. Natl. Acad. Sci. USA 91:5022-5026; Lockhart et al., 1996, Nature Biotech 14:1675; U.S. Pat. Nos. 5,578,832; 5,556,752; and 5,510,270; Blanchard et al., 1996, 11: 687-90).

Other methods for making microarrays, e.g., by masking (Maskos and Southern, 1992, Nuc. Acids Res. 20:1679-1684), may also be used. In principal, any type of array, for example, dot blots on a nylon hybridization membrane (see Sambrook et al., Molecular Cloning - A Laboratory Manual (2nd Ed.), Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1989), could be used, although, as will be recognized by those of skill in the art.

The nucleic acids to be contacted with the microarray may be prepared in a variety of ways, and may include nucleotides of the subject invention. Such nucleic acids are often labeled fluorescently. Nucleic acid hybridization and wash conditions are chosen so that the population of labeled nucleic acids will specifically hybridize to appropriate, complementary nucleic acids affixed to the matrix. Non-specific binding of the labeled nucleic acids to the array can be decreased by treating the array with a large quantity of non-specific DNA -- a so-called "blocking" step.

When fluorescently labeled probes are used, the fluorescence emissions at each site of a transcript array may be detected by scanning confocal laser microscopy. When two fluorophores are used, a separate scan, using the appropriate excitation line, is carried out for each of the two fluorophores used. Fluorescent microarray scanners are commercially available from Affymetrix, Packard BioChip Technologies, BioRobotics and many other suppliers. Signals are recorded, quantitated and analyzed using a variety of computer software.

According to the method of the invention, the relative abundance of an mRNA in two cells or cell lines is scored as a perturbation and its magnitude determined (i.e., the abundance is different in the two sources of mRNA tested), or as not perturbed (i.e., the

relative abundance is the same). As used herein, a difference between the two sources of RNA of at least a factor of about 25% (RNA from one source is 25% more abundant in one source than the other source), more usually about 50%, even more often by a factor of about 2 (twice as abundant), 3 (three times as abundant) or 5 (five times as abundant) is scored as a perturbation. Present detection methods allow reliable detection of difference of an order of about 2-fold to about 5-fold, but more sensitive methods are expected to be developed.

In addition to identifying a perturbation as positive or negative, it is advantageous to determine the magnitude of the perturbation. This can be carried out, as noted above, by calculating the ratio of the emission of the two fluorophores used for differential labeling, or by analogous methods that will be readily apparent to those of skill in the art.

In certain embodiments, the data obtained from such experiments reflects the relative expression of each gene represented in the microarray. Expression levels in different samples and conditions may now be compared using a variety of statistical methods.

12. Pharmaceutical Compositions

Pharmaceutical compositions of this invention include any modulator identified according to the present invention, or a pharmaceutically acceptable salt thereof, and a pharmaceutically acceptable carrier, adjuvant, or vehicle. The term "pharmaceutically acceptable carrier" refers to a carrier(s) that is "acceptable" in the sense of being compatible with the other ingredients of a composition and not deleterious to the recipient thereof.

Methods of making and using such pharmaceutical compositions are also included in the invention. The pharmaceutical compositions of the invention can be administered orally, parenterally, by inhalation spray, topically, rectally, nasally, buccally, vaginally, or via an implanted reservoir. The term parenteral as used herein includes subcutaneous, intracutaneous, intravenous, intramuscular, intra articular, intrasynovial, intrasternal, intrathecal, intralesional, and intracranial injection or infusion techniques.

Dosage levels of between about 0.01 and about 100 mg/kg body weight per day, preferably between about 0.5 and about 75 mg/kg body weight per day of the modulators described herein are useful for the prevention and treatment of disease and conditions, including *P. aeruginosa* mediated diseases and conditions. The amount of active ingredient that may be combined with the carrier materials to produce a single dosage form will vary depending upon the host treated and the particular mode of administration. A typical

preparation will contain from about 5% to about 95% active compound (w/w). Alternatively, such preparations contain from about 20% to about 80% active compound.

13. Antimicrobial Agents

The polypeptides of the invention may be used to develop antimicrobial agents for use in a wide variety of applications. The uses are as varied as surface disinfectants, topical pharmaceuticals, personal hygiene applications (e.g., antimicrobial soap, deodorant or the like), additives to cell culture medium, and systemic pharmaceutical products. Antimicrobial agents of the invention may be incorporated into a wide variety of products and used to treat an already existing microbial infection/contamination or may be used prophylactically to suppress future infection/contamination.

The antimicrobial agents of the invention may be administered to a site, or potential site, of infection/contamination in either a liquid or solid form. Alternatively, the agent may be applied as a coating to a surface of an object where microbial growth is undesirable using nonspecific absorption or covalent attachment. For example, implants or devices (such as linens, cloth, plastics, heart pacemakers, surgical stents, catheters, gastric tubes, endotracheal tubes, prosthetic devices) can be coated with the antimicrobials to minimize adherence or persistence of bacteria during storage and use. The antimicrobials may also be incorporated into such devices to provide slow release of the agent locally for several weeks during healing. The antimicrobial agents may also be used in association with devices such as ventilators, water reservoirs, air-conditioning units, filters, paints, or other substances. Antimicrobials of the invention may also be given orally or systemically after transplantation, bone replacement, during dental procedures, or during implantation to prevent colonization with bacteria.

In another embodiment, antimicrobial agents of the invention may be used as a food preservative or in treating food products to eliminate potential pathogens. The latter use might be targeted to the fish and poultry industries that have serious problems with enteric pathogens which cause severe human disease. In a further embodiment, the agents of the invention may be used as antimicrobials for food crops, either as agents to reduce post harvest spoilage or to enhance host resistance. The antimicrobials may also be used as preservatives in processed foods either alone or in combination with antibacterial food additives such as lysozymes.

In another embodiment, the antimicrobials of the invention may be used as an additive to culture medium to prevent or eliminate infection of cultured cells with a pathogen.

EXEMPLIFICATION

The invention now being generally described, it will be more readily understood by reference to the following examples which are included merely for purposes of illustration of certain aspects and embodiments of the present invention, and are not intended to limit the invention in any way.

EXAMPLE 1 Isolation and Cloning of Nucleic Acid

Pseudomonas aeruginosa is an opportunistic Gram-negative bacilli found in sewage, plants, and sometimes the intestine. It is capable of infecting various organs and has been identified in numerous infections including those in the ears, lungs, urinary tract, blood and in burns and surgical wound infections. Polynucleotide sequences were obtained from The Institute of Genomic Research (TIGR) (Rockville, MD; www.tigr.org). Chromosomal DNA was acquired from the American Type Culture Collection (ATCC; reference #17933D).

The coding sequences of the subject nucleic acid sequences (predicted) are obtained by reference to either publicly available databases or from the use of a bioinformatics program that is used to select the coding sequence of interest from the applicable genome. For example, coding sequences for the genome of *P. aeruginosa* may be obtained from NCBI (http://www.ncbi.nlm.nih.gov/cgi-bin/Entrez/framik?db=Genome&gi=163).

The coding DNA is amplified from purified genomic DNA using PCR with primers that are identified with a computer program. The PCR primers are selected so as to introduce restriction enzyme cleavage sites at the flanking regions of the DNA (e.g., NdeI and BglII). The forward and reverse primers have SEQ ID NO: 5 and SEQ ID NO: 6. The sequences of the primers are shown in FIGURE 5, and their respective restriction sites and melting temperatures are shown in Table 1 of FIGURE 6.

The PCR reaction is performed using 50-100 *ng* of chromosomal DNA and 2 Units of a high fidelity DNA Polymerase (for example *Pfu* Turbo (Stratagene) or Platinum *Pfx* (Invitrogen)). The thermocycling conditions for the PCR process include a DNA melting step at 94°C for 45 sec, a primer annealing step at 48°C - 58°C (depending on Primer [Tm]) for 45 sec, and an extension step at 68°C – 72°C (depending on enzyme) for 1 min 45 sec –

2 min 30 sec (depending on size of DNA). After 25-30 cycles, a final blocking step at 72°C for 9 min is carried out.

The amplified nucleic acid product is isolated from the PCR cocktail using silica-gel membrane based column chromatography (Qiagen). The quality of the PCR product is assessed by resolving an aliquot of amplified product on a 1% agarose gel. The DNA is quantified spectrophotometrically at A_{260} or by visualizing the resolved genes with a 302 nm UV-B light source.

The PCR product is directionally cloned into the polylinker region of any of three expression vectors: pET28 (Novagen), pET15 (Novagen) or pGEX (Pharmacia/LKB Biotechnology). Additional restriction enzyme sites may be engineered into the expressions vectors to allow for simultaneous clones to be prepared having different purification tags. After the ligation reaction, the DNA is transformed into competent *E. coli* cells (Strains XL1-Blue (Stratagene) or DH5 (Invitrogen)) via heat shock or electroporation as described in Sambrook, et al., Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989). The expression vectors contain the bacteriophage T7 promoter for RNA polymerase, and the *E. coli* strain used produces T7 RNA polymerase upon induction with isopropyl-β-D-thiogalactoside (IPTG). The sequence of the cloning site adds a Glutathione S-transferase (GST) tag, or a polyhistidine (6X His) tag, at the N- or C– terminus of the recombinant protein. The cloning site also inserts a cleavage site for the thrombin or Tev (Invitrogen) enzymes between the recombinant protein and the N- or C- terminal GST or polyhistidine tag.

Transformants are selected using the appropriate antibiotic (Ampicillin or Kanamycin) and identified using PCR, or another method, to analyze their DNA. The polynucleotide sequence cloned into the expression construct is then isolated using a modified alkaline lysis method (Birnboim, H.C., and Doly, J. (1979) *Nucl. Acids Res.* 7, 1513-1522.) The sequence of the clone is verified by standard polynucleotide sequencing methods. The published nucleic acid and amino acid sequences are presented in FIGURE 1 and FIGURE 2. The experimentally determined nucleic acid sequence is presented in FIGURE 3, and the amino acid sequence predicted from the sequence of FIGURE 3 is presented in FIGURE 4.

The expression construct is transformed into a bacterial host strain BL21-Gold (DE3) supplemented with a plasmid called pUBS520, which directs expression of tRNA for

arginine (agg and aga) and serves to augment the expression of the recombinant protein in the host cell (Gene, vol. 85 (1989) 109-114). The expression construct may also be transformed into BL21-Gold (DE3) without pUBS520, BL21-Gold (DE3) Codon-Plus (RIL) or (RP) (Stratagene) or Roseatta (DE3) (Novagen), the latter two of which contain genes encoding tRNAs. Alternatively, the expression construct may be transformed into BL21 STAR *E. coli* (Invitrogen) cells which has an Rnase deficiency that reduces degradation of recombinant mRNA transcript and therefore increases the protein yield. The recombinant protein is then assayed for positive overexpression in the host and the presence of the protein in the cytoplasmic (water soluble) region of the cell.

EXAMPLE 2 Test Protein Expression and Solubility

(a) Test Expression

Transformed cells are grown in LB medium supplemented with the appropriate antibiotics up to a final concentration of 100 μg/ml. The cultures are shaken at 37°C until they reach an optical density (OD_{600}) between 0.6 and 0.7. The cultures are then induced with isopropyl-beta-D-thiogalactopyranoside (IPTG) to a final concentration of 0.5 mM at 15°C for 10 hours, 25°C for 4 hours, or 30°C for 4 hours.

(b) Method One for Determining Protein Solubility Levels

The cells are harvested by centrifugation and subjected to a freeze/thaw cycle. The cells are lysed using detergent, sonication, or incubation with lysozyme. Total and soluble proteins are assayed using a 26-well BioRad Criterion gel running system. The proteins are stained with an appropriate dye (Coomassie, Silver stain, or Sypro-Red) and visualized with the appropriate visualization system. Typically, recombinant protein is seen as a prominent band in the lanes of the gel representing the soluble fraction.

(c) Method Two for Determining Protein Solubility Levels

The soluble and insoluble fractions (in the presence of 6M urea) of the cell pellet are bound to the appropriate affinity column. The purified proteins from both fractions are analysed by SDS-PAGE and the levels of protein in the soluble fraction are determined.

The approximate percent solubility of the polypeptide having the sequence of SEQ ID NO: 4 is determined using one of the foregoing methods, and the resulting percent solubility is presented in Table 1 of FIGURE 6.

EXAMPLE 3 Native Protein Expression

The expression construct clone encoding the soluble polypeptide having the amino acid sequence of SEQ ID NO: 4 is introduced into an expression host. The resultant cell line is then grown in culture. The method of growth is dependant on whether the protein to be purified is a native protein or a labeled protein. For native and ^{15}N labeled protein production, a Gold-pUBS520 (as described above), BL21-Gold (DE3) Codon-Plus (RIL) or (RP), or BL21 STAR *E. Coli* cell line is used. For generating proteins metabolically labeled with selenium, the clone is introduced into a strain called B834 (Novagen). The methods for expressing labeled polypeptides of the invention are described in the Examples that follow.

In one method for expressing an unlabeled polypeptide of the invention, 2L LB cultures or 1L TB cultures are inoculated with a 1% (v/v) starter culture (OD_{600} of 0.8). The cultures are shaken at 37°C and 200 rpm and grown to an OD_{600} of 0.6-0.8 followed by induction with 0.5mM IPTG at 15°C and 200 rpm for at least 10 hours or at 25°C for 4 hours.

The cells are harvested by centrifugation and the pellets are resuspended in 25 ml HEPES buffer (50 mM, pH 7.5), supplemented with 100µl of protease inhibitors (PMSF and benzamidine (Sigma)) and flash-frozen in liquid nitrogen.

Alternatively, for an unlabeled polypeptide of the invention, a starter culture is prepared in a 300 mL Tunair flask (Shelton Scientific) by adding 20 mL of medium having 47.6 g/L of Terrific Broth and 1.5% glycerol in dH_2O followed by autoclaving for 30 minutes at 121°C and 15 psi. When the broth cools to room temperature, the medium is supplemented with 6.3 µM $CoCl_2$-$6H_2O$, 33.2 µM $MnSO_4$-$5H_2O$, 5.9 µM $CuCl_2$-$2H_2O$, 8.1 µM H_3BO_3, 8.3 µM Na_2MoO_4-$2H_2O$, 7 µM $ZnSO_4$-$7H_2O$, 108 µM $FeSO_4$-$7H_2O$, 68 µM $CaCl_2$-$2H_2O$, 4.1 µM $AlCl_3$-$6H_2O$, 8.4 µM $NiCl_2$-$6H_2O$, 1 mM $MgSO_4$, 0.5% v/v of Kao and Michayluk vitamins mix (Sigma; Cat. No. K3129), 25 µg/mL Carbenicillin, and 50 µg/mL Kanamycin. The medium is then inoculated with several colonies of the freshly transformed expression construct of interest. The culture is incubated at 37°C and 260 rpm for about 3 hours and then transferred to a 2.5L Tunair Flask containing 1L of the above media. The 1L culture is then incubated at 37°C with shaking at 230-250 rpm on an orbital shaker having a 1 inch orbital diameter. When the culture reaches an OD_{600} of 3-6 it is induced with 0.5 mM IPTG. The induced culture is then incubated at 15°C with shaking at

230-250 rpm or faster for about 6-15 hours. The cells are harvested by centrifugation at 3500 rpm at 4°C for 20 minutes and the cell pellet is resuspended in 15 mL ice cold binding buffer (Hepes 50 mM, pH 7.5) and 100 µl of protease inhibitors (50 mM PMSF and 100 mM Benzamidine, stock concentration) and flash frozen.

EXAMPLE 4 Expression of Selmet Labeled Polypeptides

The freshly transformed cell, harboring a plasmid with a nucleic acid encoding a polypeptide of the invention, is inoculated into 20 ml of NMM (New Minimal Medium) and shaken at 37°C for 8-9 hours. This culture is then transferred into a 6L Erlenmeyer flask containing 2L of minimum medium (M9). The media is supplemented with all amino acids except methionine. All amino acids are added as a solution except for Tyrosine, Tryptophan and Phenylalanine which are added to the media in powder format. As well the media is supplemented with $MgSO_4$ (2mM final concentration), $FeSO_4.7H_2O$ (25mg/L final concentration), Glucose (0.4% final concentration), $CaCl_2$ (0.1mM final concentration) and Seleno-L-Methionine (40mg/L final concentration). When the OD_{600} of the cell culture reaches 0.8-0.9, IPTG (0.4 mM final concentration) is added to the medium for protein induction, and the cell culture is kept shaking at 15°C for 10 hours. The cells are harvested by centrifugation at 3500 rpm at 4°C for 20 minutes and the cell pellet is resuspended in 15 mL cold binding buffer (Hepes 50 mM, pH 7.5) and 100 µl of protease inhibitors (PMSF and Benzamidine) and flash frozen.

Alternatively, a starter culture is prepared in a 300 mL Tunair flask (Shelton Scientific) by adding 50 mL of sterile medium having 10% 10XM9 (37.4 mM NH_4Cl (Sigma; Cat. No. A4514), 44 mM KH_2PO_4 (Bioshop, Ontario, Canada; Cat. No. PPM 302), 96 mM Na_2HPO_4 (Sigma; Cat. No. S2429256), and 96 mM $Na_2HPO_4 7H_2O$ (Sigma; Cat. No. S9390) final concentration), 450 µM alanine, 190 µM arginine, 302 µM asparagine, 300 µM aspartic acid, 330 µM cysteine, 272 µM glutamic acid, 274 µM glutamine, 533 µM glycine, 191 µM histidine, 305 µM isoleucine, 305 µM leucine, 220 µM lysine, 242 µM phenylalanine, 348 µM proline, 380 µM serine, 336 µM threonine, 196 µM tryptophan, 220 µM tyrosine, and 342 µM valine, 204 µM Seleno-L-Methionine (Sigma; Cat. No. S3132), 0.5% v/v of Kao and Michayluk vitamins mix (Sigma; Cat. No. K3129), 2 mM $MgSO_4$ (Sigma; Cat. No. M7774), 90 µM $FeSO_4 7H_2O$ (Sigma; Cat. No. F8633), 0.4% glucose (Sigma; Cat. No. G-5400), 100 µM $CaCl_2$ (Bioshop, Ontario, Canada; Cat. No. CCL 302),

50 µg/mL Ampicillin, and 50 µg/mL Kanamycin in dH_2O. The medium is then inoculated with several colonies of *E. coli* B834 cells (Novagen) freshly transformed with an expression construct clone encoding the polypeptide of interest. The culture is then incubated at 37°C and 200 rpm until it reaches an OD_{600} of ~1 and is then transferred to a 2.5L Tunair Flask containing 1L of the above media. The 1L culture is incubated at 37°C with shaking at 200 rpm until the culture reaches an OD_{600} of 0.6-0.8 and is then induced with 0.5 mM IPTG. The induced culture is incubated overnight at 15°C with shaking at 200 rpm. The cells are harvested by centrifugation at 4200 rpm at 4°C for 20 minutes and the cell pellet is resuspended in 15 mL ice cold binding buffer (Hepes 50 mM, pH 7.5) and 100 µl of protease inhibitors (50 mM PMSF and 100 mM Benzamidine, stock concentration) and flash frozen.

Alternatively, the cell harboring a plasmid with a nucleic acid encoding a polypeptide of the invention is inoculated into 10 ml of M9 minimum medium and kept shaking at 37°C for 8-9 hours. This culture is then transferred into a 2L Baffled Flask (Corning) containing 1L minimum medium. The media is supplemented with all amino acids except methionine. All are added as a solution, except for Phenylalanine, Alanine, Valine, Leucine, Isoleucine, Proline, and Tryptophan which are added to the media in powder format. As well the media is supplemented with $MgSO_4$ (2mM final concentrtion), $FeSO_4 7H_2O$ (25 mg/L final concentration), Glucose (0.5% final concentration), $CaCl_2$ (0.1 mM final concentration) and Seleno-Methionine (50 mg/L final concentration). When the OD_{600} of the cell culture reaches 0.8-0.9, IPTG (0.8 mM final concentration) is added to the medium for protein induction, and the cell culture is kept shaking at 25°C for 4 hours. The cells are harvested by centrifuged at 3500 rpm at 4°C for 20 minutes and the cell pellet is resuspended in 10 mL cold binding buffer (Hepes 50 mM, pH 7.5) and 100 µl of protease inhibitors (PMSF and Benzamidine) and flash frozen.

EXAMPLE 5 Expression of ^{15}N Labeled Polypeptides

The cell, harboring a plasmid with a nucleic acid encoding a polypeptide of the invention, is inoculated into 2L of minimal media (containing ^{15}N isotope, Cambridge Isotope Lab) in a 6L Erlenmeyer flask. The minimal media is supplemented with 0.01 mM $ZnSO_4$, 0.1 mM $CaCl_2$, 1 mM $MgSO_4$, 5 mg/L Thiamine.HCl, and 0.4% glucose. The 2L culture is grown at 37°C and 200 rpm to an OD_{600} of between 0.7-0.8. The culture is then

induced with 0.5 mM IPTG and allowed to shake at 15°C for 14 hours. The cells are harvested by centrifugation and the cell pellet is resuspended in 15 mL cold binding buffer and 100µl of protease inhibitor and flash frozen. The protein is then purified as described below.

Alternatively, the freshly transformed cell, harboring a plasmid with the gene of interest, is inoculated into 10 mL of M9 media (with ^{15}N isotope) and supplemented with with 0.01 mM $ZnSO_4$, 0.1 mM $CaCl_2$, 1 mM $MgSO_4$, 5 mg/L Thiamine.HCl, and 0.4% glucose. After 8-10 hours of growth at 37°C, the culture is transferred to a 2L Baffled flask (Corning) containing 990 mL of the same media. When OD_{600} of the culture is between 0.7-0.8, protein production is initiated by adding IPTG to a final concentration of 0.8 mM and lowering the temperature to 25°C. After 4 hours of incubation at this temperature, the cells are harvested, and the cell pellet is resuspended in 10 mL cold binding buffer (Hepes 50 mM, pH 7.5) and 100 µl of protease inhibitor and flash frozen.

EXAMPLE 6 Method One for Purifying Polypeptides of the Invention

The frozen pellets are thawed and sonicated to lyse the cells (5 x 30 seconds, output 4 to 5, 80% duty cycle, in a Branson Sonifier, VWR). The lysates are clarified by centrifugation at 14,000 rpm for 60 min at 4°C to remove insoluble cellular debris. The supernatants are removed and supplemented with 1 µl of Benzonase Nuclease (25 U/µl, Novagen).

The recombinant protein is purified using DE52 (anion exchanger, Whatman) and Ni-NTA columns (Qiagen). The DE52 columns (30 mm wide, Biorad) are prepared by mixing 10 grams of DE52 resin in 25 ml of 2.5 M NaCl per protein sample, applying the resin to the column and equilibrating with 30 ml of binding buffer (50 mM in HEPES, pH 7.5, 5% glycerol (v/v), 0.5 M NaCl, 5 mM imidazole). Ni-NTA columns are prepared by adding 3.5-8 ml of resin to the column (20 mm wide, Biorad) based on the level of expression of the recombinant protein and equilibrating the column with 30 ml of binding buffer. The columns are arranged in tandem so that the protein sample is first passed over the DE52 column and then loaded directly onto the Ni-NTA column.

The Ni-NTA columns are washed with at least 150 ml of wash buffer (50mM HEPES, pH 7.5, 5% glycerol (v/v), 0.5 M NaCl, 30 mM imidazole) per column. A pump may be used to load and/or wash the columns. The protein is eluted off of the Ni-NTA

column using elution buffer (50 mM in HEPES, pH 7.5, 5% glycerol (v/v), 0.5 M NaCl, 250 mM imidazole) until no more protein is observed in the aliquots of eluate as measured using Bradford reagent (Biorad). The eluate is supplemented with 1 mM of EDTA and 0.2 mM DTT.

The samples are assayed by SDS-PAGE and stained with Coomassie Blue, with protein purity determined by visual staining.

Two methods may be used to remove the His tag located at either the C or N-terminus. In certain instances, the His tag may not be removed. In either case, the expressed polypeptide will have additional residues attributable to the His tag, as shown in the following table:

SEQ ID NO for Additional Residues	*Additional Residues*	*Type of Tag and Whether or Not Removed*
N/A	GSH	His tag removed from N-terminus
SEQ ID NO: 7	MGSSHHHHHHSSGLVPRGSH	His tag not removed from N-terminus
SEQ ID NO: 8	GSENLYFQGHHHHHH	His tag removed from C-terminus
SEQ ID NO: 9	GSENLYFQ	His tag not removed from C-terminus

In method one, a sample of purified polypeptide is supplemented with 2.5 mM $CaCl_2$ and an appropriate amount of thrombin (the amount added will vary depending on the activity of the enzyme preparation) and incubated for ~20-30 minutes on ice in order to remove the His tag. In method two, a sample of purified polypeptide is combined with thirty units of recombinant TEV protease in 50 mmol TRIS HCl pH = 8.0, 0.5 mmol EDTA and 1 mmol DTT, followed by incubation at 4°C overnight, to remove the His tag.

The protein sample is then dialyzed in dialysis buffer (10mM HEPES, pH 7.5, 5% glycerol (v/v) and 0.5 M NaCl) for at least 8 hours using a Slide-A-Lyzer (Pierce) appropriate for the molecular weight of the recombinant protein. An aliquot of the cleaved and dialyzed samples is then assayed by SDS-PAGE and stained with Coomassie Blue to determine the purity of the protein and the success of cleavage.

The remainder of the sample is centrifuged at 2700 rpm at 4°C for 10-15 minutes to remove any precipitant and supplemented with 100 µl of protease inhibitor cocktail (0.1 M benzamidine and 0.05 M PMSF) (NO Bioshop). The protein is then applied to a second Ni-NTA column (~8 ml of resin) to remove the His-tags and eluted with binding buffer or

wash buffer until no more protein is eluting off the column as assayed using the Bradford reagent. The eluted sample is supplemented with 1 mM EDTA and 0.6 mM of DTT and concentrated to a final volume of ~15 mls using a Millipore Concentrator with an appropriately sized filter at 2700 rpm at 4°C. The samples are then dialyzed overnight against crystallization buffer and concentrated to final volume of 0.3-0.7 ml.

EXAMPLE 7 Method Two for Purifying Polypeptides of the Invention

The frozen pellets are thawed and supplemented with 100 µl of protease inhibitor (0.1 M benzamidine and 0.05 M PMSF), 0.5% CHAPS, and 4 U/ml Benzonase Nuclease. The sample is then gently rocked on a Nutator (VWR, setting 3) at room temperature for 30 minutes. The cells are then lysed by sonication (1 x 30 seconds, output 4 to 5, 80% duty cycle, in a Branson Sonifier, VWR) and an aliquot is saved for a gel sample.

The recombinant protein is purified using a three column system. The columns are set up in tandem so that the lysate flows from a Biorad Econo (5.0 x 30 cm x 589 ml) "lysate" column onto a Biorad Econo (2.5 x 20 cm x 98 ml) DE52 column and finally onto a Biorad Econo (1.5 x 15 cm x 27 ml) Ni-NTA column. The lysate is mixed with 10 g of equilibrated DE52 resin and diluted to a total volume of 300 ml with binding buffer. This mixture is poured into the first column which is empty. The remainder of the purification procedure is described in EXAMPLE 6 above.

EXAMPLE 8 Method Three for Purifying Polypeptides of the Invention

The frozen pellets are thawed and sonicated to lyse the cells (5 x 30 seconds, output 4 to 5, 80% duty cycle, in a Branson Sonifier, VWR). The lysates are clarified by centrifugation at 14000 rpm for 60 min at 4°C to remove insoluble cellular debris. The supernatants are removed and supplemented with 1 µl of Benzonase Nuclease (25 U/µl, Novagen).

The recombinant protein is purified using DE52 (anion exchanger, Whatman) and Glutathione sepharose columns (Glutathione-Superflow resin, Clontech). The DE52 columns (30 mm wide, Biorad) are prepared by mixing 10 grams of DE52 resin in 20 ml of 2.5 M NaCl per protein sample, applying the resin to the column and equilibrating with 30 ml of loading buffer (50mM in HEPES, pH 7.5, 10% glycerol (v/v), 0.5 M NaCl, 1 mM EDTA, 1 mM DTT). Glutathione sepharose columns are prepared by adding 3 ml of resin to the column (20 mm wide, Biorad) and equilibrating the column with 30 ml of loading

buffer. The columns are arranged in tandem so that the protein sample is first passed over the DE52 column and then loads directly onto the Glutathione sepharose column.

The columns are washed with at least 150 ml of loading buffer supplemented with protease inhibitor cocktail (0.1 M benzamidine and 0.05 M PMSF) per column. A pump may be used to load and/or wash the columns. The protein is eluted off of the Glutathione sepharose column using elution buffer (20mM HEPES, pH 7.5, 0.5 M NaCl, 1 mM EDTA, 1 mM DTT; 25 mM glutathione (reduced form)) until no more protein is observed in the aliquots of eluate as measured using Biorad Bradford reagent.

The GST tag may be removed using thrombin or other procedures known in the art. The protein samples are then dialyzed into crystallization buffer (10 mM Hepes, pH 7.5, 500 mM NaCl) to remove free glutathione and assayed by SDS-PAGE followed by staining with Coomassie blue. Prior to use or storage, the samples are concentrated to final volume of 0.3-0.5 ml.

Using one or more of the methods described above, purified polypeptide having SEQ ID NO: 4 is obtained in a yield of approximately 3.9 mg per liter of culture. The purified polypeptide is essentially the only protein visualized in the SDS-PAGE assay using Coomassie Blue described above, which is at least about 95% or greater purity. The polypeptide so expressed and purified has the following additional amino acid residues from the removed His tag at the N-terminus: GSH.

The protein samples so prepared and purified may be used in the biophysical studies that follow, with or without the His tag or the residual amino acids resulting from removal of the His tag. In certain instances, such as EXAMPLE 11, the polypeptide used may be a fusion protein with a specific tag.

A stable solution of purified polypeptide having SEQ ID NO: 4, prepared and purified as described above, may be prepared with 11.0 mg (or a lesser amount) of protein in one ml of either the dialysis or crystallization buffers (or possibly both) described above in EXAMPLE 6 or EXAMPLE 8, respectively.

Certain of the foregoing information is also set forth in Table 1 of FIGURE 6

EXAMPLE 9 Mass Spectrometry Analysis via Fingerprint Mapping

A gel slice from a purification protocol described above containing a polypeptide of the invention is cut into 1 mm cubes and 10 to 20 µl of 1% acetic acid is added. After washing with 100 - 150 µl HPLC grade water and removal of the liquid, acetonitrile (~200

µl, approximately 3 to 4 times the volume of the gel particles) is added followed by incubation at room temperature for 10 to 15 minutes with vortexing. A second acetonitrile wash may be required to completely dehydrate the gel particles. The protein in the gel particles is reduced at 50 degrees Celsius using 10 mM dithiothreitol (in 100 mM ammonium bicarbonate) and then alkylated at room temperature in the dark using 55 mM iodoacetamide (in 100 mM ammonium bicarbonate). The gel particles are rinsed with a minimal volume of 100 mM ammonium bicarbonate before a trypsin (50 mM ammonium bicarbonate, 5 mM $CaCl_2$, and 12.5 ng/µl trypsin) solution is added. The gel particles are left on ice for 30 to 45 minutes (after 20 minutes incubation more trypsin solution is added). The excess trypsin solution is removed and 10 to 15 µl digestion buffer without trypsin is added to ensure the gel particles remain hydrated during digestion. After digestion at 37°C, the supernatant is removed from the gel particles. The peptides are extracted from the gel particles with 2 changes of 100 µL of 100 mM ammonium bicarbonate with shaking for 45 minutes and pooled with the initial gel supernatant. The extracts are acidified to 1% (v/v) with 100% acetic acid.

The tryptic peptides are purified with a C18 reverse phase resin. 250 µL of dry resin is washed twice with methanol and twice with 75% acetonitrile/1% acetic acid. A 5:1 slurry of solvent:resin is prepared with 75% acetonitrile/1% acetic acid. To the extracted peptides, 2 µL of the resin slurry is added and the solution is shaken for 30 minutes at room temperature. The supernatant is removed and replaced with 200 µL of 2% acetonitrile/1% acetic acid and shaken for 5-15 minutes. The supernatant is removed and the peptides are eluted from the resin with 15 µL of 75% acetonitrile/1% acetic acid with shaking for about 5 minutes. The peptide and slurry mixture is applied to a filter plate and centrifuged, and the filtrate is collected and stored at –70°C until use.

Alternatively, the tryptic peptides are purified using $ZipTip_{C18}$ (Millipore, Cat # ZTC18S960). The ZipTips are first pre-wetted by aspirating and dispensing 100% methanol. The tips are then washed with 2% acetonitrile/1% acetic acid (5 times), followed by 65% acetonitrile/1% acetic (5 times) and returned to 2% acetonitrile/1% acetic acid (10 times). The digested peptides are bound to the ZipTips by aspirating and dispensing the samples 5 times. Salts are removed by washing ZipTips with 2% acetonitrile/1% acetic acid (5 times). 10 µL of 65% acetonitrile/1% acetic acid is collected by the ZipTips and dispensed into a 96-well microtitre plate.

Analytical samples containing tryptic peptides are subjected to MALDI-TOF mass spectrometry. Samples are mixed 1:1 with a matrix of α-cyano-4-hydroxy-*trans*-cinnamic acid. The sample/matrix mixture is spotted on to the MALDI sample plate with a robot, either a Gilson 215 liquid handler or BioMek FX laboratory automation workstation (Beckman). The sample/matrix mixture is allowed to dry on the plate and is then introduced into the mass spectrometer. Analysis of the peptides in the mass spectrometer is conducted using both delayed extraction mode (400 ns delay) and an ion reflector to ensure high resolution of the peptides.

Internally-calibrated tryptic peptide masses are searched against databases using a correlative mass matching algorithm. The Proteometrics software package (ProteoMetrics) is utilized for batch database searching of tryptic peptide mass spectra. Statistical analysis is performed on each protein match to determine its validity. Typical search constraints include error tolerances within 0.1 Da for monoisotopic peptide masses, carboxyamidomethylation of cysteines, no oxidation of methionines allowed, and 0 or 1 missed enzyme cleavages. The software calculates the probability that a candidate in the database search is the protein being analyzed, which is expressed as the Z-score. The Z-score is the distance to the population mean in unit of standard deviation and corresponds to the percentile of the search in the random match population. If a search is in the 95th percentile, for example, about 5% of random matches could yield a higher Z-score than the search. A Z-score of 1.282 for a search indicates that the search is in the 90th percentile, a Z-score of 1.645 indicates that the search is in the 95th percentile, a Z-score of 2.326 indicates that the search is in the 99th percentile, and a Z-score of 3.090 indicates that the search is in the 99.9th percentile.

EXAMPLE 10 Mass Spectrometry Analysis via High Mass

A matrix solution of 25 mg/mL of 3,5-dimethoxy-4-hydroxycinnamic acid (sinapinic acid) in 66% (v/v) acetonitrile/1% (v/v) acetic acid is prepared along with an internal calibrant of carbonic anhydrase. On to a stainless steel polished MALDI target, 1.5 μL of a protein solution (concentration of 2 μg/μL) is spotted, followed immediately by 1.5 μL of matrix. 3 μL of 40% (v/v) acetonitrile/1% (v/v) acetic acid is then added to each spot has dried. The sample is either spotted manually or utilizing a Gilson 215 liquid handler or BioMek FX laboratory automation workstation (Beckman). The MALDI-TOF instrument

utilizes positive ion and linear detection modes. Spectra are acquired automatically over a mass to charge range from 0-150,000 Da, pulsed ion extraction delay is set at 200 ns, and 600 summed shots of 50-shot steps are completed.

The theoretical molecular weight of the protein for MALDI-TOF is determined from its amino acid sequence, taking into account any purification tag or residue thereof still present and any labels (e.g., selenomethionine or ^{15}N). To account for ^{15}N incorporation, an amount equal to the theoretical molecular weight of the protein divided by 70 is added. The mass of water is subtracted from the overall molecular weight. The MALDI-TOF spectrum is calibrated with the internal calibrant of carbonic anhydrase (observed as either $[MH^{+}_{avg}]$ 29025 or $[MH_2^{2+}]$ 14513).

EXAMPLE 11 Method One for Isolating and Identifying Interacting Proteins

(a) Method One for Preparation of Affinity Column

Micro-columns are prepared using forceps to bend the ends of P200 pipette tips and adding 10 µl of glass beads to act as a column frit. Six micro-columns are required for every polypeptide to be studied. The micro-columns are placed in a 96-well plate that has 1 mL wells. Next, a series of solutions of the polypetide having SEQ ID NO: 4 or other polypeptide of the invention, prepared and purified as described above and with a GST tag on either terminus, is prepared so as to give final amounts of 0, 0.1, 0.5, 1.0, and 2.0 mg of ligand per ml of resin volume.

A slurry of Glutathione-Sepharose 4B (Amersham) is prepared and 0.5 ml slurry/ligand is removed (enough for six 40-µg aliquots of resin). Using a glass frit Buchner funnel, the resin is washed sequentially with three 10 ml portions each of distilled H_2O and 1 M ACB (20 mM HEPES pH 7.9, 1 M NaCl, 10% glycerol, 1 mM DTT, and 1 mM EDTA). The Glutathione-Sepharose 4B is completely drained of buffer, but not dried. The Glutathione-Sepharose 4B is resuspended as a 50% slurry in 1 M ACB and 80 µl is added to each micro-column to obtain 40 µg/column. The buffer containing the ligand concentration series is added to the columns and allowed to flow by gravity. The resin and ligand are allowed to cross-link overnight at 4°C. In the morning, micro-columns are washed with 100 µl of 1 M ACB and allowed to flow by gravity. This is repeated twice more and the elutions are tested for cross-linking efficiency by measuring the amount of

unbound ligand. After washing, the micro-columns are equilibrated using 200 µl of 0.1 M ACB (20 mM HEPES pH 7.5, 0.1 M NaCl, 10% glycerol, 1 mM DTT, 1 mM EDTA).

In another method, the recombinant GST fusion protein can be replaced by a hexa-histidine fusion peptide for use with NTA-Agarose (Qiagen) as the solid support. No adaptation to the above protocol is required for the substitution of NTA agarose for GST Sepharose except that the recombinant protein requires a six histidine fusion peptide in place of the GST fusion.

(b) Method Two for Preparation of Affinity Column

In an alternative method, GST-Sepharose 4B may be replaced by Affi-gel 10 Gel (Bio-Rad). The column resin for affinity chromatography could also be Affigel 10 resin which allows for covalent attachment of the protein ligand to the micro affinity column. An adaptation to the above protocol for the use of this resin is a pre-wash of the resin with 100% isopropanol. No fusion peptides or proteins are required for the use of Affigel 10 resin.

(c) Method One for Bacterial Extract Preparation

A *P. aeruginosa* extract is prepared from cell pellets using a French press followed by sonication. An *P. aeruginosa* cell pellet (~6 g) is suspended in 3 pellet volumes (~20 ml final volume) of 20 mM HEPES pH 7.5, 150 mM NaCl, 10% glycerol, 10 mM $MgSO_4$, 10 mM $CaCl_2$, 1 mM DTT, 1 mM PMSF, 1 mM benzamidine, 40 µg/ml RNAse A, 75 units/ml S1 nuclease, and 40 units/ml DNAse 1. The cell suspension is lysed with one pass with a French Pressure Cell followed by sonication on ice using three bursts of 20 seconds each. The lysate is agitated at 4°C for 30 minutes, brought up to 0.5 M NaCl and then incubated for an additional 30 min at 4°C with agitation. The lysate is centrifuged at 25,000 rpm for 1 hr at 4°C in a Ti70 fixed angle Beckman rotor. The supernatant is removed and dialyzed overnight in a 10,000 Mr dialysis membrane against dialysis buffer (20 mM HEPES pH 7.5, 10 % glycerol, 1 mM DTT, 1 mM EDTA, 100 mM NaCl, 10 mM $MgSO_4$, 10 mM $CaCl_2$, 1 mM benzamidine, and 1 mM PMSF). The dialyzed protein extract is removed from the dialysis tubing and frozen in one ml aliquots at –70°C.

(d) Method Two for Bacterial Extract Preparation

Bacterial cell extracts from *P. aeruginosa* are prepared from cell pellets using a Bead-Beater apparatus (Bio-spec Products Inc.) and zirconia beads (0.1 mm diameter). The bacterial cell pellet is suspended (~6 g) is suspended in 3 pellet volumes (~20 ml final

volume) of 20 mM HEPES pH 7.5, 150 mM NaCl, 10% glycerol, 10 mM $MgSO_4$, 10 mM $CaCl_2$, 1 mM DTT, 1 mM PMSF, 1 mM benzamidine, 40 µg/ml RNAse A, 75 units/ml S1 nuclease, and 40 units/ml DNAse 1. The cells are lysed with 10 pulses of 30 sec between 90 sec pauses at a temperature of -5 °C. The lysate is separated from the zirconia beads using a standard column apparatus. The lysate is centrifuged at 20000 rpm (48000 x g) in a Beckman JA25.50 rotor. The supernatant is removed and dialyzed overnight at 4 °C against dialysis buffer (20 mM HEPES pH 7.5, 10 % glycerol, 1 mM DTT, 1 mM EDTA, 100 mM NaCl, 10 mM $MgSO_4$, 10 mM $CaCl_2$, 1 mM benzamidine, and 1 mM PMSF). The dialyzed protein extract is removed from the dialysis tubing and frozen in one ml aliquots at –70°C.

(e) HeLa Cell Extract Preparation

A HeLa cell extract is prepared in the presence of protease inhibitors. Approximately 30 g of Hela cells are submitted to a freeze/thaw cycle and then divided into two tubes. To each tube 20 ml of Buffer A (10 mM HEPES pH 7.9, 1.5 mM MgCl, 10 mM KCl, 0.5 mM DTT, 0.5 mM PMSF) and a protease inhibitor cocktail are added. The cell suspension is homogenized with 10 strokes (2 x 5 strokes) to lyse the cells. Buffer B (15 ml per tube) is added (50 mM HEPES pH 7.9, 1.5 mM MgCl, 1.26 M NaCl, 0.5 mM DTT, 0.5 mM PMSF, 0.5 mM EDTA, 75% glycerol) to each tube followed by a second round of homogenization (2 x 5 strokes). The lysates are stirred on ice for 30 minutes followed by centrifugation 37,000 rpm for 3 hr at 4°C in a Ti70 fixed angle Beckman rotor. The supernatant is removed and dialyzed overnight in a 10,000 Mr dialysis membrane against dialysis buffer (20 mM HEPES pH 7.9, 10% glycerol, 1 mM DTT, 1 mM EDTA, and 1 M NaCl. The dialyzed protein extract is removed from the dialysis tubing and frozen in one ml aliquots at –70°C.

(f) Affinity Chromatography

Cell extract is thawed and diluted to 5 mg/ml prior to loading 5 column volumes onto each micro-column. Each column is washed with 5 column volumes of 0.1 M ACB. This washing is repeated once. Each column is then washed with 5 column volumes of 0.1 M ACB containing 0.1% Triton X-100. The columns are eluted with 4 column volumes of 1% sodium dodecyl sulfate into a 96 well PCR plate. To each eluted fraction is added one-tenth volume of 10-fold concentrated loading buffer for SDS-PAGE.

(g) Resolution of the Eluted Proteins and Detection of Bound Proteins

The components of the eluted samples are resolved on SDS-polyacrylamide gels containing 13.8% polyacrylamide using the Laemmli buffer system and stained with silver

nitrate. The bands containing the interacting protein are excised with a clean scalpel. The gel volume is kept to a minimum by cutting as close to the band as possible. The gel slice is placed into one well of a low protein binding, 96-well round-bottom plate. To the gel slices is added 20 µl of 1% acetic acid.

EXAMPLE 12 Method Two for Isolating and Identifying Interacting Proteins

Interacting proteins may be isolated using immunoprecipitation. Naturally-occurring bacterial or eukaryotic cells are grown in defined growth conditions or the cells can be genetically manipulated with a protein expression vector. The protein expression vector is used to transiently transfect the cDNA of interest into eukaryotic or prokaryotic cells and the protein is expressed for up to 24 or 48 hours. The cells are harvested and washed three times in sterile 20 mM HEPES (pH7.4)/Hanks balanced salts solution (H/H). The cells are finally resuspended in culture media and incubated at 37°C for 4-8 hr.

The harvested cells may be subjected to one or more culture conditions that may alter the protein profile of the cells for a given period of time. The cells are collected and washed with ice-cold H/H that includes 10 mM sodium pyrophosphate, 10 mM sodium fluoride, 10 mM EDTA, and 1 mM sodium orthovanadate. The cells are then lysed in lysis buffer (50 mM Tris-HCl (pH 8.0), 150 mM NaCl, 1% Triton X-100, 10 mM sodium pyrophosphate, 10mM sodium fluoride, 10 mM EDTA, 1 mM sodium orthovanadate, 1 µg/mL PMSF, 1 µg/mL aprotinin, 1 µg/mL leupeptin, and 1 µg/mL pepstatin A) by gentle mixing, and placed on ice for 5 minutes. After lysis, the lysate is transferred to centrifuge tubes and centrifuged in an ultracentrifuge at 75000 rpm for 15 min at 4°C. The supernatant is transferred to eppendorf tubes and pre-cleared with 10 µl of rabbit pre-immune antibody on a rotator at 4°C for 1 hr. Forty µl of protein A-Sepharose (Amersham) is then added and incubated at 4°C overnight on a rotator.

The protein A-Sepharose beads are harvested and the supernatant removed to a fresh eppendorf tube. Immune antibody is added to supernatant and rotated for 1 hr at 4°C. Thirty µl of protein A-Sepharose is then added and the mixture is further rotated at 4°C for 1 hr. The beads are harvested and the supernatant is aspirated. The beads are washed three times with 50 mM Tris (pH 8.0), 150 mM NaCl, 0.1% Triton X-100, 10 mM sodium fluoride, 10 mM sodium pyrophosphate, 10 mM sodium orthovanadate, and 10 mM EDTA. Dry the beads with a 50 µl Hamilton syringe. Laemmli loading buffer containing 100 mM

DTT is added to the beads and samples are boiled for 5 min. The beads are spun down and the supernatant is loaded onto SDS-PAGE gels. Comparison of the control and experimental samples allows for the selection of polypeptides that interact with the protein of interest.

EXAMPLE 13 Sample for Mass Spectrometry of Interacting Proteins

The gel slices are cut into 1 mm cubes and 10 to 20 µl of 1% acetic acid is added. The gel particles are washed with 100 - 150 µl of HPLC grade water (5 minutes with occasional mixing), briefly centrifuged, and the liquid is removed. Acetonitrile (~200 µl, approximately 3 to 4 times the volume of the gel particles) is added followed by incubation at room temperature for 10 to 15 minutes with vortexing. A second acetonitrile wash may be required to completely dehydrate the gel particles. The sample is briefly centrifuged and all the liquid is removed.

The protein in the gel particles is reduced at 50 degrees Celsius using 10 mM dithiothreitol (in 100 mM ammonium bicarbonate) for 30 minutes and then alkylated at room temperature in the dark using 55 mM iodoacetamide (in 100 mM ammonium bicarbonate). The gel particles are rinsed with a minimal volume of 100 mM ammonium bicarbonate before a trypsin (50 mM ammonium bicarbonate, 5 mM $CaCl_2$, and 12.5 ng/µl trypsin) solution is added. The gel particles are left on ice for 30 to 45 minutes (after 20 minutes incubation more trypsin solution is added). The excess trypsin solution is removed and 10 to 15 µl digestion buffer without trypsin is added to ensure the gel particles remain hydrated during digestion. The samples are digested overnight at 37°C.

The following day, the supernatant is removed from the gel particles. The peptides are extracted from the gel particles with 2 changes of 100 µL of 100 mM ammonium bicarbonate with shaking for 45 minutes and pooled with the initial gel supernatant. The extracts are acidified to 1% (v/v) with 100% acetic acid.

(a) Method One for Purification of Tryptic Peptides

The tryptic peptides are purified with a C18 reverse phase resin. 250 µL of dry resin is washed twice with methanol and twice with 75% acetonitrile/1% acetic acid. A 5:1 slurry of solvent : resin is prepared with 75% acetonitrile/1% acetic acid. To the extracted peptides, 2 µL of the resin slurry is added and the solution is shaken at moderate speed for 30 minutes at room temperature. The supernatant is removed and replaced with 200 µL of

2% acetonitrile/1% acetic acid and shaken for 5-15 minutes with moderate speed. The supernatant is removed and the peptides are eluted from the resin with 15 µL of 75% acetonitrile/1% acetic acid with shaking for about 5 minutes. The peptide and slurry mixture is applied to a filter plate and centrifuged for 1-2 minutes at 1000 rpm, the filtrate is collected and stored at –70°C until use.

(b) Method Two for Purification of Tryptic Peptides

Alternatively, the tryptic peptides may be purified using ZipTip$_{C18}$ (Millipore, Cat # ZTC18S960). The ZipTips are first pre-wetted by aspirating and dispensing 100% methanol 5 times. The tips are then washed with 2% acetonitrile/1% acetic acid (5 times), followed by 65% acetonitrile/1% acetic (5 times) and returned to 2% acetonitrile/1% acetic acid (5 times). The ZipTips are replaced in their rack and the residual solvent is eliminated. The ZipTips are washed again with 2% acetonitrile/1% acetic acid (5 times). The digested peptides are bound to the ZipTips by aspirating and dispensing the samples 5 times. Salts are removed by washing ZipTips with 2% acetonitrile/1% acetic acid (5 times). 10 µL of 65% acetonitrile/1% acetic acid is collected by the ZipTips and dispensed into a 96-well microtitire plate. 1 µL of sample and 1 µL of matrix are spotted on a MALDI-TOF sample plate for analysis.

EXAMPLE 14 Mass Spectrometric Analysis of Interacting Proteins

(a) Method One for Analysis of Tryptic Peptides

Analytical samples containing tryptic peptides are subjected to Matrix Assisted Laser Desorption/Ionization Time Of Flight (MALDI-TOF) mass spectrometry. Samples are mixed 1:1 with a matrix of α-cyano-4-hydroxy-*trans*-cinnamic acid. The sample/matrix mixture is spotted on to the MALDI sample plate with a robot. The sample/matrix mixture is allowed to dry on the plate and is then introduced into the mass spectrometer. Analysis of the peptides in the mass spectrometer is conducted using both delayed extraction mode and an ion reflector to ensure high resolution of the peptides.

Internally-calibrated tryptic peptide masses are searched against both in-house proprietary and public databases using a correlative mass matching algorithm. Statistical analysis is performed on each protein match to determine its validity. Typical search constraints include error tolerances within 0.1 Da for monoisotopic peptide masses and

carboxyamidomethylation of cysteines. Identified proteins are stored automatically in a relational database with software links to SDS-PAGE images and ligand sequences.

(b) Method Two for Analysis of Tryptic Peptides

Alternatively, samples containing tryptic peptides are analyzed with an ion trap instrument. The peptide extracts are first dried down to approximately 1 μL of liquid. To this, 0.1% trifluoroacetic acid (TFA) is added to make a total volume of approximately 5 μL. Approximately 1-2 μL of sample are injected onto a capillary column (C8, 150 μm ID, 15 cm long) and run at a flow rate of 800 nL/min. using the following gradient program:

Time (minutes)	% Solvent A	% Solvent B
0	95	5
30	65	35
40	20	80
41	95	5

Where Solvent A is composed of water/0.5% acetic acid and Solvent B is acetonitrile/0.5% acetic acid. The majority of the peptides will elute between the 20-40 % acetonitrile gradient. Two types of data from the eluting HPLC peaks are acquired with the ion trap mass spectrometer. In the MS^1 dimension, the mass to charge range for scanning is set at 400-1400 - this will determine the parent ion spectrum. Secondly, the instrument has MS^2 capabilities whereby it will acquire fragmentation spectra of any parent ions whose intensities are detected to be greater than a predetermined threshold (Mann and Wilm, *Anal Chem* 66(24): 4390-4399 (1994)). A significant amount of information is collected for each protein sample as both a parent ion spectrum and many daughter ion spectra are generated with this instrumentation.

All resulting mass spectra are submitted to a database search algorithm for protein identification. A correlative mass algorithm is utilized along with a statistical verification of each match to identify a protein's identification (Ducret A, et al., *Protein Sci* 7(3): 706-719 (1998)). This method proves much more robust than MALDI-TOF mass spectrometry for identifying the components of complex mixtures of proteins.

No interacting proteins were observed using at least one of the methods described above.

EXAMPLE 15 NMR Analysis

Purified protein sample is centrifuged at 13,000 rpm for 10 minutes with a bench-top microcentrifuge to eliminate any precipitated protein. The supernatant is then transferred into a clean tube and the sample volume is measured. If the sample volume is less than 450 µl, an appropriate amount of crystal buffer is added to the sample to reach that volume. Then 50 µl of D_2O (99.9%) is added to the sample to make an NMR sample of 500 µl. The usual concentration of the protein sample is usually approximately 1 mmol or greater.

NMR screening experiments are performed on a Bruker AV600 spectrometer equipped with a cryoprobe, or other equivalent instrumentation. All spectra are recorded at 25°C. Standard 1D proton pulse sequence with presaturation is used for 1D screening. Normally, a sweepwidth of 6400 Hz, and eight or sixteen scans are used, although different pulse sequences are known to those of skill in the art and may be readily determined. For ^{1}H, ^{15}N HSQC experiments, a pulse sequence with "flip-back" water suppression may be used. Typically, sweepwidths of 8000 Hz and 2000 Hz are used for F2 and F1 dimension, respectively. Four to sixteen scans are normally adequate. The data is then processed on a Sun Ultra 5 computer with NMRpipe software.

EXAMPLE 16 X-ray Crystallography

(a) Crystallization

Subsequent to purification, a subject polypeptide is centrifuged for 10 minutes at 4°C and at 14,000 rpm in order to sediment any aggregated protein. The protein sample is then diluted in order to provide multiple concentrations for screening.

Two 96 well plates (Nunc) are employed for the initial crystal screen, with 48 potential crystallization conditions. The screening library has crystallization conditions found in Hampton Research Crystal Screen I (Jankarik, J. and S.H. Kim, J. Appl. Cryst., 1991. 24:409-11), Hampton Research Crystal Screen II, Hampton Crystal Screen I-Lite, and from Emerald Biostructures, Inc., Bainbridge Island, WA, Wizard I, Wizard II, Cryo I and Cryo II. Alternatively, other conditions known to those of skill in the art, including those provided in screening kits available from other companies, may also be tested.

Conditions are tested at multiple protein concentrations and at two temperatures (4 and 20°C). Crystal setups may be performed by a liquid handling robot appropriately

programmed for sitting drop experiments. The robot loads 50 µl of buffer into each screening well on a 24 or 96 well sitting drop crystal screen tray, and then loads 1 - 5 µl of protein into each drop reservoir to be screened on the plate. Subsequently, the robot loads 1.5 µl of the corresponding screening solution into the drop reservoir atop the protein. The plate is then sealed using transparent tape, and stored at 4 or 20°C. Each plate is observed two days, two weeks, and 1 month after being set. Alternatively, screens may be performed using 0.1 - 10 µl drops suspended at the interface of two immiscible oils. The protein containing solution has a density intermediate between the two oils and thus floats between them (Chayen N.E.: 1996, *Protein Eng.* 9:927–29). This procedure may be performed in an automated fashion by an appropriately programmed liquid handling robot, with additional steps being required initially to introduce the oils. No tape is added to facilitate gradual drying out of the drop to promote crystallization.

Having identified conditions that are best suited for further crystal refinement, subsequent plates are set up to explore the affects of variables such as temperature, pH, salt or PEG concentration on crystal size and form, with the intent of establishing conditions where the protein is able to form crystals of suitable size and morphology for diffraction analysis. Each refinement is performed in the sitting drop format in a 24 well Lindbro plate. Each well in the tray contains 500 µl of screening solution, and a 1.5 µl drop of protein diluted with 1.5 µl of the screening solution is set to hang from the siliconized glass cover slip covering the well. Alternatively, refinement steps may be performed using either the machine 96 well plate hanging drop method or the oil suspension method described above.

Suitable crystals for x-ray experimentation were obtained by using 11 mg/ml protein 20% PEG4000 as precipitant, 6.4% isopropanol, 2% MPD and 100mM HEPES, pH 7.4. The crystals are somewhat sensitive to changes in their solution and care is required to freeze them without the crystals cracking. The crystal was stabilized in a mixture of 20% (v/v) ethylene glycol in the mother liquor. A slow soak protocol was used wherein ethylene glycol and the mother liquor were mixed in a 1:3 volume ratio, and 5 microlitres were added on top of the skin that covers the surface of the drop containing the crystals. This skin was then punctured at a location remote from the crystal, and the two liquids allowed to mix. After a few minutes of passive mixing, a loop was used to ensure that the liquids

were thoroughly mixed. The crystals were frozen at 100K in a cold gas stream generated by an X-stream 2000 cryocooling device.

(b) Co-Crystallization

A variety of methods known in the art may be used for preparation of co-crystals comprising the subject polypeptides and one or more compounds that interact with the subject polypeptides, such as, for example, an inhibitor, co-factor, substrate, polynucleotide, polypeptide, and/or other molecule. In one exemplary method, crystals of the subject polypeptide may be soaked, for an appropriate period of time, in a solution containing a compound that interacts with a subject polypeptide. In another method, solutions of the subject polypeptide and/or compound that interacts with the subject polypeptide may be prepared for crystallization as described above and mixed into the above-described sitting drops. In certain embodiments, the molecule to be co-crystallized with the subject polypeptide may be present in the buffer in the sitting drop prior to addition of the solution comprising the subject polypeptide. In other embodiments, the subject polypeptide may be mixed with another molecule before adding the mixture to the sitting drop. Based on the teachings herein, one of skill in the art may determine the co-crystallization method yielding a co-crystal comprising the subject polypeptide.

(c) Heavy Atom Substitution

For preparation of crystals containing heavy atoms, crystals of the subject polypeptide may be soaked in a solution of a compound containing the appropriate heavy atom for such period as time as may be experimentally determined is necessary to obtain a useful heavy atom derivative for x-ray purposes. Likewise, for other compounds that may be of interest, including, for example, inhibitors or other molecules that interact with the subject polypeptide, crystals of the subject polypeptide may be soaked in a solution of such compound for an appropriate period of time.

(d) Data collection and processing

Native data were collected from a single frozen crystal using a monochromatic CuK x-rays (Bruker X-ray generator) and a SMART 6000 CCD detector. The crystal diffracted to a minimum d-spacing of 1.9 Å. All data were processed using the SAINT/PROSCALE software package. The protein atom coordinates of the *E. coli* PTH (Protein data bank code 2PTH) were used as the search model for molecular replacement using CNX. This yielded a solution with a correlation coefficient of 50.5 and a packing coefficient of 67.9. The structure was refined using the CNX program suite. Ten percent

of the reflections were randomly excluded from the refinement, and used to monitor Rfree. A maximum likelihood target (with a flat bulk solvent correction and no low resolution or sigma cutoff applied to the data) was used in the refinement protocol. Refinement of the model using the simulated annealing slow-cooling and individual temperature factor refinement protocol was alternated with manual inspection and rebuilding of the model using TURBO-FRODO (Roussel A. et Cambillau C. (1989) TURBO-FRODO. Silicon Graphics Geometry Partner Directory, Silicon Graphics, Mountain view, California). After several cycles of refinement and manual rebuilding, 192 residues of the protein monomer were modeled, with residues 193 – 194 being disordered. In addition to the protein, the structure also contains 215 solvent molecules and one ethylene glycol molecule.

Structure solution and refined statistics for are reported in Table 3, contained in FIGURE 8. FIGURE 9 contains a list of the atomic coordinates of the subject polypeptide and other molecules contained in the crystal. FIGURE 10 to FIGURE 14 depict various features of the crystal structure and other properties of a subject polypeptide.

(e) Analysis of the X-ray Structure of the Subject Polypeptide

General Description of Structure

Although the crystal used contains two molecules in the asymmetric unit, PTH behaves in solution as a monomer of 193 residues. The R.M.S deviation between the two monomers of the asymmetric unit is 0.23 Å for 193 C atoms showing that the two monomers are essentially identical. The PA enzyme structure is composed of a single / globular domain build around a twisted -sheet (FIGURE 11).

Active Site and Other Druggable Regions

The amino acid sequences for the druggable regions that are two active binding pockets of the enzyme revealed by the *P. aeruginosa* structure using the Site Finder functionality in a software called MOE (Chemical Computing Group) are indicated below:

Binding Site Identifier	*Amino Acid Residues*	*Description of Binding Site*
Binding Site I	N12, Y17, H22, Y68, M69, N70, D95, L97, G114, H115, N116, G117, V147, S148, V151, L152	*Peptide binding/catalytic site.* This pocket is approximately 105 units (Å^3).
Binding Site II	A11, N12, P13, G14, P15, Y17, D18, N19,	*tRNA binding site.* This pocket is approximately 171

	T20, R21, N23, A24, A26, R27, L49, I64, P65, T66, G153, R154, A155, Q160, L163	units ($Å^3$).

In the *E. coli* enzyme, the peptide binding site pocket extends for 255 units, compared to 105 units ($Å^3$) in the *P. aeruginosa*. Furthermore, the tRNA binding site in *E. coli* appears to be split into two (85 and 87 units) due to the presence of a tryptophan residue in the narrow region of the channel (FIGURE 14). The specific presence of a bulky residue in the *E. coli* PTH (FIGURE 14) partially occluding the mid-section of the active site channel suggests that conformational changes may be required upon tRNA-PTH complex formation to accommodate tRNAs carrying longer peptides.

The putative catalytic sites of the *P. aeruginosa* enzyme located on both edges of the protein dimer are highly conserved to the ones found in the *E. coli* monomeric protein (FIGURE 13).

Comparison to Other Peptidyl-tRNA Hydrolase Enzymes

The superimposition of the *P. aeruginosa* and *E. coli* PTHs produces a final r.m.s. deviation of 1.1 Å for 188 pairs of C compared (FIGURE 12). Residues shown to involved in catalysis or tRNA substrate binding are extremely well conserved among the five bacteria (alignment shown in FIGURE 10 with the exception of K142, shown to be involved in the tRNA binding in the *E. coli* PTH that is part of a region making a flexible lid over the active site crevice. The conservation as determined from this alignment for PTH residues involved in catalysis or tRNA binding are shown below:

Conservation of PTH residues shown to be involved in catalysis or tRNA binding

Residues	*P. aeruginosa*	*E. coli*	*H. pylori*	*S. aureus*	*S. pneumonia*
Asn12	Conserved	Conserved	Conserved	Conserved	Conserved
His22	Conserved	Conserved	Conserved	Conserved	Conserved
Met69	Conserved	Conserved	Conserved	Conserved	Conserved
Asn70	Conserved	Conserved	Conserved	Conserved	Conserved
Asp95	Conserved	Conserved	Conserved	Conserved	Conserved
His115	Conserved	Conserved	Conserved	Conserved	Conserved

Asn116	Conserved	Conserved	Conserved	Conserved	Conserved
Arg135	Conserved	Conserved	Conserved	Conserved	Conserved
Lys144	Ser	Lys	Deletion	Gly	Gly

Based in part on the structural information described above, in one aspect, the present invention is directed to a druggable region of a polypeptide of the invention containing a druggable region for binding a polypeptide with a volume of less than approximately 200, 150 or 125 units (Å^3). In one embodiment, the druggable region comprises at least one of residues Asn12, Tyr17, His22, Tyr68, Met69, Asn70, Asp95, Leu97, Gly114, His115, Asn116, Gly117, Val147, Ser148, Val151, or Leu152. In another embodiment, the druggable region comprises at least one of residues Ala11, Asn12, Pro13, Gly14, Pro15, Tyr17, Asp18, Asn19, Thr20, Arg21, Asn23, Ala24, Ala26, Arg27, Leu49, Ile64, Pro65, Thr66, Gly153, Arg154, Ala155, Gln160, or Leu163. In still another embodiment, the druggable region may comprise at least one of residues Asn12, His22, Met69, Asn70, Asp95, His115, Asn116, Arg135, or Lys144. In another aspect, the present invention is directed to a druggable region of a peptidyl-tRNA hydrolase for binding tRNA that (i) does not contain any amino acid residue, such as tryptophan or an amino acid residue containing an aromatic group or a large aliphatic nonpolar group, or (ii) that consists of a single pocket with a volume of at least 100 or 150 units (Å^3). In another aspect, the present invention is directed to a peptidyl-tRNA hydrolase that does not require a conformational change to form a complex with tRNA. In another aspect, the present invention is directed to a peptidyl-tRNA hydrolase that is a dimer of two monomers in its naturally occurring active form.. In another aspect, the present invention is directed toward an inhibitor that interacts with the active site of such an enzyme. In one embodiment, an inhibitor may modulate the conformation change in a peptidyl-tRNA hydrolase.

EXAMPLE 17 Annotations

The functional annotation is arrived at by comparing the amino acid sequence of the ORF against all available ORFs in the NCBI database using BLAST. The closest match is selected to provide the probable function of the polypeptide having the sequence of SEQ ID NO: 2. Results of this comparison are described above and set forth in Table 2 of FIGURE 7.

The COGs database (Tatusov RL, Koonin EV, Lipman DJ. Science 1997; 278 (5338) 631-37) classifies proteins encoded in twenty-one completed genomes on the basis of sequence similarity. Members of the same Cluster of Orthologous Group, ("COG"), are expected to have the same or similar domain architecture and the same or substantially similar biological activity. The database may be used to predict the function of uncharacterised proteins through their homology to characterized proteins. The COGs database may be searched from NCBI's website (http://www.ncbi.nlm.nih.gov/COG/) to determine functional annotation descriptions, such as "information storage and processing" (translation, ribosomal structure and biogenesis, transcription, DNA replication, recombination and repair); "cellular processes" (cell division and chromosome partitioning, post-translational modification, protein turnover, chaperones, cell envelope biogenesis, outer membrane, cell motility and secretion, inorganic ion transport and metabolism, signal transduction mechanisms);' or "metabolism" (energy production and conversion, carbohydrate transport and metabolism, amino acid transport and metabolism, nucleotide transport and metabolism, coenzyme metabolism, lipid metabolism). For certain polypeptides, there is no entry available. Results of this analysis are described above and set forth in Table 2 of FIGURE 7.

EXAMPLE 18 Essential Gene Analysis

SEQ ID NO: 2 is compared to a number of publicly available "essential genes" lists to determine whether that protein is encoded by an essential gene. An example of such a list is descended from a free release at the www.shigen.nig.ac.jp PEC (profiling of *E. coli* chromosome) site, http://www.shigen.nig.ac.jp/ecoli/pec/. The list is prepared as follows: a wildcard search for all genes in class "essential" yields the list of essential *E. coli* proteins encoded by essential genes, which number 230. These 230 hits are pruned by comparing against an NCBI *E. coli* genome. Only 216 of the 230 genes on the list are found in the NCBI genome. These 216 are termed the essential-216-ecoli list. The essential-216-ecoli list is used to garner "essential" genes lists for other microbial genomes by blasting. For instance, formatting the 216-ecoli as a BLAST database, then BLASTing a genome (e.g. *S. aureus*) against it, elucidates all *S. aureus* genes with significant homology to a gene in the 216-essential list. SEQ ID NO: 2 is compared against the appropriate list and a match with a score of e^{-25} or better is considered an essential gene according to that list. In addition to the list described above, other lists of essential genes are publicly available or may be

determined by methods disclosed publicly, and such lists and methods are considered in deciding whether a gene is essential. See, for example, Thanassi et al., Nucleic Acids Res 2002 Jul 15;30(14):3152-62; Forsyth et al., Mol Microbiol 2002 Mar;43(6):1387-400; Ji et al., Science 2001 Sep 21;293(5538):2266-9; Sassetti et al., Proc Natl Acad Sci U S A 2001 Oct 23;98(22):12712-7; Reich et al., J Bacteriol 1999 Aug;181(16):4961-8; Akerley et al., Proc Natl Acad Sci U S A 2002 Jan 22;99(2):966-71). Also, other methods are known in the art for determing whether a gene is essential, such as that disclosed in U.S. Patent Application No. 10/202,442 (filed July 24, 2002). The conclusion as to whether the gene encoding the amino acid sequence set forth in SEQ ID NO: 2 is essential is set forth in Table 2 of FIGURE 7.

EXAMPLE 19 PDB Analysis

SEQ ID NO: 2 is compared against the amino acid sequences in a database of proteins whose structures have been solved and released to the PDB (protein data bank). The identity/information about the top PDB homolog (most similar "hit", if any; a PDB entry is only considered a hit if the score is e^{-4} or better) is annotated, and the percent similarity and identity between SEQ ID NO: 2 and the closest hit is calculated, with both being indicated in Table 2 of FIGURE 7.

EXAMPLE 20 Virtual Genome Analysis

VGDB or VG is a queryable collection of microbial genome databases annotated with biophysical and protein information. The organisms present in VG include:

File	*GRAM*	*Species*	*Source*	*Genome file date*
ecoli.faa	G-	*Escherichia coli*	NCBI	November 18 1998
hpyl.faa	G-	*Helicobacter pylori*	NCBI	April 19 1999
paer.faa	G-	*Pseudomonas aeruginosa*	NCBI	September 22 2000
ctra.faa	G-	*Chlamydia trachomatis*	NCBI	December 22 1999
hinf.faa	G-	*Haemophilus influenzae*	NCBI	November 26 1999
nmen.faa	G-	*Neisseria meningitidis*	NCBI	December 28 2000
rpxx.faa	G-	*Rickettsia prowazekii*	NCBI	December 22 1999
bbur.faa	G-	*Borrelia burgdorferi*	NCBI	November 11 1998
bsub.faa	G+	*Bacillus subtilis*	NCBI	December 1 1999
staph.faa	G+	*Staphylococcus aureus*	TIGR	March 8 2001
spne.faa	G+	*Streptococcus pneumoniae*	TIGR	February 22 2001
mgen.faa	G+	*Mycoplasma genitalium*	NCBI	November 23 1999

efae.faa G+ *Enterococcus faecalis* TIGR March 8 2001

The VGDB comprises 13 microbial genomes, annotated with biophysical information (pI, MW, etc), and a wealth of other information. These 13 organism genomes are stored in a single flatfile (the VGDB) against which PSI-blast queries can be done.

SEQ ID NO: 2 is queried against the VGDB to determine whether this sequence is found, conserved, in many microbial genomes. There are certain criteria that must be met for a positive hit to be returned (beyond the criteria inherent in a basic PSI-blast).

When an ORF is queried it may have a maximum of 13 VG-organism hits. A hit is classified as such as long as it matches the following criteria: Minimum Length (as percentage of query length): 75 *(Ensure hit protein is at least 75% as long as query)*; Maximum Length (as percentage of query length): 125 *(Ensure hit protein is no more than 125% as long as query)*; eVal:-10 *(Ensure hit has an e-Value of e-10 or better)*; Id%:>:25 *(Ensure hit protein has at least 25% identity to query)*. The e-Value is a standard parameter of BLAST sequence comparisons, and represents a measure of the similarity between two sequences based on the likelihood that any similarities between the two sequences could have occurred by random chance alone. The lower the e-Value, the less likely that the similarities could have occurred randomly and, generally, the more similar the two sequences are.

The organisms having an orthologue of the polypeptide having SEQ ID NO: 2 are listed in Table 2, shown in FIGURE 7.

EXAMPLE 21 Epitopic Regions

The three most likely epitopic regions of a polypeptide having SEQ ID NO: 2 are predicted using the semi-empirical method of Kolaskar and Tongaonkar (FEBS Letters 1990 v276 172-174), the software package called Protean (DNASTAR), or MacVectors's Protein analysis tools (Accerlyrs). The antigenic propensity of each amino acid is calculated by the ratio between frequency of occurrence of amino acids in 169 antigenic determinants experimentally determined and the calculated frequency of occurrence of amino acids at the surface of protein. The results of these bioinformatics analyses are presented in Table 2, shown in FIGURE 7.

EQUIVALENTS

The present invention provides among other things, novel proteins, protein structures and protein-protein interactions. While specific embodiments of the subject invention have been discussed, the above specification is illustrative and not restrictive. Many variations of the invention will become apparent to those skilled in the art upon review of this specification. The full scope of the invention should be determined by reference to the claims, along with their full scope of equivalents, and the specification, along with such variations.

All publications and patents mentioned herein, including those items listed below, are hereby incorporated by reference in their entirety as if each individual publication or patent was specifically and individually indicated to be incorporated by reference. In case of conflict, the present application, including any definitions herein, will control. To the extent that any U.S. Provisional Patent Applications to which this patent application claims priority incorporate by reference another U.S. Provisional Patent Application, such other U.S. Provisional Patent Application is not incorporated by reference herein unless this patent application expressly incorporates by reference, or claims priorty to, such other U.S. Provisional Patent Application.

Also incorporated by reference in their entirety are any polynucleotide and polypeptide sequences which reference an accession number correlating to an entry in a public database, such as those maintained by The Institute for Genomic Research (TIGR) (www.tigr.org) and/or the National Center for Biotechnology Information (NCBI) (www.ncbi.nlm.nih.gov).

Also incorporated by reference are the following: WO 00/45168, WO 00/79238, WO 00/77712, EP 1047108, EP 1047107, WO 00/72004, WO 00/73787, WO00/67017, WO 00/48004, WO 01/48209, WO 00/45168, WO 00/45164, U.S.S.N. 09/720272; PCT/CA99/00640; U.S. Patent Application Nos: 10/097125 (filed March 12, 2002); 10/097193 (filed March 12, 2002); 10/202442 (filed July 24, 2002); 10/097194 (filed March 12, 2002); 09/671817 (filed September 17, 2000); 09/965654 (filed September 27, 2001); 09/727812 (filed November 30, 2000); 60/370667 (filed April 8, 2002); a utility patent application entited "Methods and Appartuses for Purification" (filed September 18, 2002); U.S. Patent Numbers 6451591; 6254833; 6232114; 6229603; 6221612; 6214563; 6200762; 6171780; 6143492; 6124128; 6107477; D428157; 6063338; 6004808; 5985214; 5981200; 5928888; 5910287; 6248550; 6232114; 6229603; 6221612; 6214563; 6200762;

6197928; 6180411; 6171780; 6150176; 6140132; 6124128; 6107066; 6270988; 6077707; 6066476; 6063338; 6054321; 6054271; 6046925; 6031094; 6008378; 5998204; 5981200; 5955604; 5955453; 5948906; 5932474; 5925558; 5912137; 5910287; 5866548; 6214602; 5834436; 5777079; 5741657; 5693521; 5661035; 5625048; 5602258; 5552555; 5439797; 5374710; 5296703; 5283433; 5141627; 5134232; 5049673; 4806604; 4689432; 4603209; 6217873; 6174530; 6168784; 6271037; 6228654; 6184344; 6040133; 5910437; 5891993; 5854389; 5792664; 6248558; 6341256; 5854922; 5866343; 6,274,361; 6,303,771; Jorgensen & Kurkland, J Mol Biol 1990, 215, 511-21; Vogel Z, Zamir A & Elison D, PNAS 1968, 61, 701-7; Heurgue-Hamard V, Mora L, Guarneros G, Buckhingham RH, EMBO J 1996, 15, 2826-2833; Bonin, P. D., G. H. Choi, et al. (2002) Protein Expr Purif 24(1): 123-30; Valadez, J. G., J. Hernandez-Sanchez, et al. (2001) Mol Microbiol 39(2): 361-9; Menez, J., V. Heurgue-Hamard, et al. (2000) Nucleic Acids Res 28(23): 4725-32; Thanedar, S., N. V. Kumar, et al. (2000) J Biol Chem 275(27): 20361-7; Fromant, M., P. Plateau, et al. (2000) Biochemistry 39(14): 4062-7; Cruz-Vera, L. R., I. Toledo, et al. (2000) J Bacteriol 182(6): 1523-8; Dincbas, V., V. Heurgue-Hamard, et al. (1999) J Mol Biol 291(4): 745-59; Fromant, M., P. Plateau, et al. (1999) Biochemistry 38(16): 4982-7; Tenson, T., J. V. Herrera, et al. (1999) J Bacteriol 181(5): 1617-22; Lluisma, A. O. and M. A. Ragan (1998) Curr Genet 34(2): 112-9; Karimi, R., M. Y. Pavlov, et al. (1998) J Mol Biol 281(2): 241-52; Heurgue-Hamard, V., R. Karimi, et al. (1998) Embo J 17(3): 808-16; Hernandez, J., C. Ontiveros, et al. (1997) Biochimie 79(8): 527-31; Schmitt, E., Y. Mechulam, et al. (1997) Embo J 16(15): 4760-9; Ontiveros, C., J. G. Valadez, et al. (1997) J Mol Biol 269(2): 167-75; Schmitt, E., M. Fromant, et al. (1997) Proteins 28(1): 135-6; De La Vega, F. M., J. M. Galindo, et al. (1996) Gene 169(1): 97-100; Schmitt, E., J. M. Guillon, et al. (1996) Biochimie 78(7): 543-54; Ouzounis, C., P. Bork, et al. (1995) Protein Sci 4(11): 2424-8; Menninger, J. R. (1995) J Basic Clin Physiol Pharmacol 6(3-4): 229-50; Galindo, J. M., G. Guarneros, et al. (1994) Gene 151(1-2): 153-6; Menninger, J. R. and R. A. Coleman (1993) Antimicrob Agents Chemother 37(9): 2027-9; Dutka, S., T. Meinnel, et al. (1993) Nucleic Acids Res 21(17): 4025-30; Meinnel, T., Y. Mechulam, et al. (1993) Biochimie 75(12): 1061-75; Aumercier, M., S. Bouhallab, et al. (1992) J Antimicrob Chemother 30(Suppl A): 9-14; Gross, M., P. Crow, et al. (1992) J Biol Chem 267(3): 2080-6; Gross, M., T. K. Starn, et al. (1992) J Biol Chem 267(3): 2073-9; Murgola, E. J. and G. Guarneros (1991) Biochimie 73(12): 1573-8; Garcia-Villegas, M. R., F. M. De La Vega, et al. (1991) Embo J 10(11): 3549-55; Gotz, F., C. Fleischer, et al. (1989) Eur J Biochem

183(1): 19-24; Petrullo, L. A. and D. Elseviers (1986) J Bacteriol 165(2): 608-11; Hove-Jensen, B. (1985) Mol Gen Genet 201(2): 269-76; Caplan, A. B. and J. R. Menninger (1984) Mol Gen Genet 194(3): 534-8; Menninger, J. R., A. B. Caplan, et al. (1983) Mol Gen Genet 190(2): 215-21; Menninger, J. R. and D. P. Otto (1982) Antimicrob Agents Chemother 21(5): 811-8; Haenni, A. L., S. Joshi, et al. (1982) Prog Nucleic Acid Res Mol Biol 27: 85-104; Abraham, A. K. and A. Pihl (1980) Eur J Biochem 106(1): 257-62; Schimmel, P. R. (1980) CRC Crit Rev Biochem 9(3): 207-51; Atherly, A. G. (1979) Mol Gen Genet 175(3): 305-11; Vidales, F. J., C. Bernabeu, et al. (1979) Biochemistry 18(19): 4155-8; Menninger, J. R. (1979) J Bacteriol 137(1): 694-6; Menninger, J. R. (1976) J Biol Chem 251(11): 3392-8; Ganoza, M. C., N. Barraclough, et al. (1976) Eur J Biochem 65(2): 613-22; Johnson, A. E., W. R. Woodward, et al. (1976) Biochemistry 15(3): 569-75; Shiloach, J., S. Bauer, et al. (1975) Isr J Med Sci 11(11): 1195-6; Shiloach, J., Y. Lapidot, et al. (1975) FEBS Lett 57(2): 130-3; Zucker, W. V. (1975) Anal Biochem 63(2): 522-7; Schulman, L. H. and H. Pelka (1975) J Biol Chem 250(2): 542-7; Menninger, J. R., C. Walker, et al. (1973) Mol Gen Genet 121(4): 307-24; Vogel, Z., T. Vogel, et al. (1971) Eur J Biochem 21(4): 582-92; Brun, G., D. Paulin, et al. (1971) Biochimie 53(2): 225-31; Menninger, J. R., M. C. Mulholland, et al. (1970) Biochim Biophys Acta 217(2): 496-511; Kossel, H. (1970) Biochim Biophys Acta 204(1): 191-202; De Groot, N., Y. Groner, et al. (1969) Biochim Biophys Acta 186(2): 286-96; Chen, et al. (2000) Biotechniques 28(3): 498-500, 504-5; Stover, et al. (2000) Nature 406(6799): 959-964; McLaughlin, et al. (1993) J. Biol. Chem. 268(9): 6470-6; and Bonin, et al. (2002) Anal. Biochem. 306(1):8-16.

We claim:

CLAIMS

1. A composition comprising an isolated, recombinant polypeptide, wherein the polypeptide comprises: (a) an amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; (b) an amino acid sequence having at least about 95% identity with the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; or (c) an amino acid sequence encoded by a polynucleotide that hybridizes under stringent conditions to the complementary strand of a polynucleotide having SEQ ID NO: 1 or SEQ ID NO: 3 and has at least one biological activity of peptidyl-tRNA hydrolase from *P. aeruginosa*; and wherein the polypeptide of (a), (b) or (c) is at least about 90% pure in a sample of the composition.

2. The composition of claim 1, wherein the polypeptide is at least about 95% pure as determined by gel electrophoresis.

3. The composition of claim 1, wherein the polypeptide is purified to essential homogeneity.

4. The composition of claim 1, wherein at least about two-thirds of the polypeptide in the sample is soluble.

5. The composition of claim 1, wherein the polypeptide is fused to at least one heterologous polypeptide that increases the solubility or stability of the polypeptide.

6. The composition of claim 1, which further comprises a matrix suitable for mass spectrometry.

7. The composition of claim 6, wherein the matrix is a nicotinic acid derivative or a cinnamic acid derivative.

8. A sample comprising an isolated, recombinant polypeptide, wherein the polypeptide comprises: (a) an amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; (b) an amino acid sequence having at least about 95% identity with the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; or (c) an amino acid sequence encoded by a polynucleotide that hybridizes under stringent conditions to the complementary strand of a polynucleotide having SEQ ID NO: 1 or SEQ ID NO: 3 and has at least one biological activity of peptidyl-tRNA hydrolase from *P. aeruginosa*; and wherein the polypeptide of (a), (b) or (c) is labeled with a heavy atom.

9. The sample of claim 8, wherein the heavy atom is one of the following: cobalt, selenium, krypton, bromine, strontium, molybdenum, ruthenium, rhodium, palladium,

silver, cadmium, tin, iodine, xenon, barium, lanthanum, cerium, praseodymium, neodymium, samarium, europium, gadolinium, terbium, dysprosium, holmium, erbium, thulium, ytterbium, lutetium, tantalum, tungsten, rhenium, osmium, iridium, platinum, gold, mercury, thallium, lead, thorium and uranium.

10. The sample of claim 8, wherein the polypeptide is labeled with seleno-methionine.

11. The sample of claim 8, further comprising a cryo-protectant.

12. The sample of claim 11, wherein the cryo-protectant is one of the following: methyl pentanediol, isopropanol, ethylene glycol, glycerol, formate, citrate, mineral oil and a low-molecular-weight polyethylene glycol.

13. A crystallized, recombinant polypeptide comprising: (a) an amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; (b) an amino acid sequence having at least about 95% identity with the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; or (c) an amino acid sequence encoded by a polynucleotide that hybridizes under stringent conditions to the complementary strand of a polynucleotide having SEQ ID NO: 1 or SEQ ID NO: 3 and has at least one biological activity of peptidyl-tRNA hydrolase from *P. aeruginosa*; wherein the polypeptide of (a), (b) or (c) is in crystal form.

14. A crystallized complex comprising the crystallized, recombinant polypeptide of claim 13 and a co-factor, wherein the complex is in crystal form.

15. A crystallized complex comprising the crystallized, recombinant polypeptide of claim 13 and a small organic molecule, wherein the complex is in crystal form.

16. The crystallized, recombinant polypeptide of claim 13, which diffracts x-rays to a resolution of about 3.5 Å or better.

17. The crystallized, recombinant polypeptide of claim 13, wherein the polypeptide comprises at least one heavy atom label.

18. The crystallized, recombinant polypeptide of claim 17, wherein the polypeptide is labeled with seleno-methionine.

19. A method for designing a modulator for the prevention or treatment of *P. aeruginosa* related disease or disorder, comprising:

(a) providing a three-dimensional structure for a crystallized, recombinant polypeptide of claim 13;

(b) identifying a potential modulator for the prevention or treatment of *P. aeruginosa* related disease or disorder by reference to the three-dimensional structure;

(c) contacting a polypeptide of the composition of claim 1 or *P. aeruginosa* with the potential modulator; and

(d) assaying the activity of the polypeptide or determining the viability of *P. aeruginosa* after contact with the modulator, wherein a change in the activity of the polypeptide or the viability of *P. aeruginosa* indicates that the modulator may be useful for prevention or treatment of a *P. aeruginosa* related disease or disorder.

20. A sample comprising an isolated, recombinant polypeptide, wherein the polypeptide comprises: (a) an amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; (b) an amino acid sequence having at least about 95% identity with the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; or (c) an amino acid sequence encoded by a polynucleotide that hybridizes under stringent conditions to the complementary strand of a polynucleotide having SEQ ID NO: 1 or SEQ ID NO: 3 and has at least one biological activity of peptidyl-tRNA hydrolase from *P. aeruginosa*; and wherein the polypeptide of (a), (b) or (c) is enriched in at least one NMR isotope.

21. The sample of claim 20, wherein the NMR isotope is one of the following: hydrogen-1 (^{1}H), hydrogen-2 (^{2}H), hydrogen-3 (^{3}H), phosphorous-31 (^{31}P), sodium-23 (^{23}Na), nitrogen-14 (^{14}N), nitrogen-15 (^{15}N), carbon-13 (^{13}C) and fluorine-19 (^{19}F).

22. The sample of claim 20, further comprising a deuterium lock solvent.

23. The sample of claim 22, wherein the deuterium lock solvent is one of the following: acetone (CD_3COCD_3), chloroform ($CDCl_3$), dichloro methane (CD_2Cl_2), methylnitrile (CD_3CN), benzene (C_6D_6), water (D_2O), diethylether ($(CD_3CD_2)_2O$), dimethylether ($(CD_3)_2O$), N,N-dimethylformamide ($(CD_3)_2NCDO$), dimethyl sulfoxide (CD_3SOCD_3), ethanol (CD_3CD_2OD), methanol (CD_3OD), tetrahydrofuran (C_4D_8O), toluene ($C_6D_5CD_3$), pyridine (C_5D_5N) and cyclohexane (C_6H_{12}).

24. The sample of claim 20, which is contained within an NMR tube.

25. A method for identifying small molecules that bind to a polypeptide of the composition of claim 1, comprising:

(a) generating a first NMR spectrum of an isotopically labeled polypeptide of the composition of claim 1;

(b) exposing the polypeptide to one or more small molecules;

(c) generating a second NMR spectrum of the polypeptide which has been exposed to one or more small molecules; and

(d) comparing the first and second spectra to determine differences between the first and the second spectra, wherein the differences are indicative of one or more small molecules that have bound to the polypeptide.

26. A host cell comprising a nucleic acid encoding a polypeptide comprising: (a) an amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; (b) an amino acid sequence having at least about 95% identity with the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; or (c) an amino acid sequence encoded by a polynucleotide that hybridizes under stringent conditions to the complementary strand of a polynucleotide having SEQ ID NO: 1 or SEQ ID NO: 3 and has at least one biological activity of peptidyl-tRNA hydrolase from *P. aeruginosa*; wherein a culture of the host cell produces at least about 1 mg of the polypeptide per liter of culture and the polypeptide is at least about one-third soluble as measured by gel electrophoresis.

27. An isolated, recombinant polypeptide, comprising: (a) an amino acid sequence having at least about 90% identity with the amino acid sequence set forth in SEQ ID NO: 4; or (b) an amino acid sequence encoded by a polynucleotide that hybridizes under stringent conditions to the complementary strand of a polynucleotide having SEQ ID NO: 1 or SEQ ID NO: 3 and has at least one biological activity of peptidyl-tRNA hydrolase from *P. aeruginosa*; and wherein the polypeptide comprises one or more of the following amino acid residues at the specified position of the polypeptide: Asn12, His22, Met69, Asn70, Asp95, His115, Asn116, and Arg135.

28. A method for obtaining structural information of a crystallized polypeptide, the method comprising:

(a) crystallizing a recombinant polypeptide, wherein the polypeptide comprises: (1) an amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; (2) an amino acid sequence having at least about 95% identity with the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; or (3) an amino acid sequence encoded by a polynucleotide that hybridizes under stringent conditions to the complementary strand of a polynucleotide having SEQ ID NO: 1 or SEQ ID NO: 3 and has at least one biological activity of peptidyl-tRNA hydrolase from *P. aeruginosa*; and wherein the crystallized polypeptide is capable of diffracting X-rays to a resolution of 3.5 Å or better; and

(b) analyzing the crystallized polypeptide by X-ray diffraction to determine the three-dimensional structure of at least a portion of the crystallized polypeptide.

29. The method of claim 28, wherein the three-dimensional structure of the portion of the crystallized polypeptide is determined to a resolution of 3.5 Å or better.

30. A method for identifying a druggable region of a polypeptide, the method comprising:

(a) obtaining crystals of a polypeptide comprising (1) an amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; (2) an amino acid sequence having at least about 95% identity with the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; or (3) an amino acid sequence encoded by a polynucleotide that hybridizes under stringent conditions to the complementary strand of a polynucleotide having SEQ ID NO: 1 or SEQ ID NO: 3 and has at least one biological activity of peptidyl-tRNA hydrolase from *P. aeruginosa*, such that the three dimensional structure of the crystallized polypeptide may be determined to a resolution of 3.5 Å or better;

(b) determining the three dimensional structure of the crystallized polypeptide using X-ray diffraction; and

(c) identifying a druggable region of the crystallized polypeptide based on the three-dimensional structure of the crystallized polypeptide.

31. The method of claim 30, wherein the druggable region is an active site.

32. The method of claim 31, wherein the druggable region is on the surface of the polypeptide.

33. Crystalline peptidyl-tRNA hydrolase from *P. aeruginosa* comprising a crystal having unit cell dimensions $a = 63.702$ Å, $b = 63.702$ Å, $c = 154.844$ Å, $\alpha = \beta = 90°$, $\gamma = 120$, with an hexagonal space group $P6_122$ and 2 molecules per asymmetric unit.

34. A crystallized polypeptide comprising (1) an amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; (2) an amino acid sequence having at least about 95% identity with the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; or (3) an amino acid sequence encoded by a polynucleotide that hybridizes under stringent conditions to the complementary strand of a polynucleotide having SEQ ID NO: 1 or SEQ ID NO: 3 and has at least one biological activity of peptidyl-tRNA hydrolase from *P. aeruginosa*; wherein the crystal has a $P6_122$ space group.

35. A crystallized polypeptide comprising a structure of a polypeptide that is defined by a substantial portion of the atomic coordinates set forth in FIGURE 9.

36. A method for determining the crystal structure of a homolog of a polypeptide, the method comprising:

(a) providing the three dimensional structure of a first crystallized polypeptide comprising (1) an amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; (2) an amino acid sequence having at least about 95% identity with the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; or (3) an amino acid sequence encoded by a polynucleotide that hybridizes under stringent conditions to the complementary strand of a polynucleotide having SEQ ID NO: 1 or SEQ ID NO: 3 and has at least one biological activity of peptidyl-tRNA hydrolase from *P. aeruginosa*;

(b) obtaining crystals of a second polypeptide comprising an amino acid sequence that is at least 70% identical to the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4, such that the three dimensional structure of the second crystallized polypeptide may be determined to a resolution of 3.5 Å or better; and

(c) determining the three dimensional structure of the second crystallized polypeptide by x-ray crystallography based on the atomic coordinates of the three dimensional structure provided in step (a).

37. The method of claim 36, wherein the atomic coordinates for the second crystallized polypeptide have a root mean square deviation from the backbone atoms of the first polypeptide of not more than 1.0 Å for all backbone atoms shared in common with the first polypeptide and the second polypeptide.

38. A method for homology modeling a homolog of peptidyl-tRNA hydrolase from *P. aeruginosa*, comprising:

(a) aligning the amino acid sequence of a homolog of peptidyl-tRNA hydrolase from *P. aeruginosa* with an amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 4 and incorporating the sequence of the homolog of peptidyl-tRNA hydrolase from *P. aeruginosa* into a model of peptidyl-tRNA hydrolase from *P. aeruginosa* derived from structure coordinates as listed in FIGURE 9 to yield a preliminary model of the homolog of peptidyl-tRNA hydrolase from *P. aeruginosa*;

(b) subjecting the preliminary model to energy minimization to yield an energy minimized model;

(c) remodeling regions of the energy minimized model where stereochemistry restraints are violated to yield a final model of the homolog of peptidyl-tRNA hydrolase from *P. aeruginosa*.

39. A method for obtaining structural information about a molecule or a molecular complex of unknown structure comprising:

(a) crystallizing the molecule or molecular complex;

(b) generating an x-ray diffraction pattern from the crystallized molecule or molecular complex;

(c) applying at least a portion of the structure coordinates set forth in FIGURE 9 to the x-ray diffraction pattern to generate a three-dimensional electron density map of at least a portion of the molecule or molecular complex whose structure is unknown.

40. A method for attempting to make a crystallized complex comprising a polypeptide and a modulator having a molecular weight of less than 5 kDa, the method comprising:

(a) crystallizing a polypeptide comprising (1) an amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; (2) an amino acid sequence having at least about 95% identity with the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; or (3) an amino acid sequence encoded by a polynucleotide that hybridizes under stringent conditions to the complementary strand of a polynucleotide having SEQ ID NO: 1 or SEQ ID NO: 3 and has at least one biological activity of peptidyl-tRNA hydrolase from *P. aeruginosa*; such that crystals of the crystallized polypeptide will diffract x-rays to a resolution of 5 Å or better; and

(b) soaking the crystals in a solution comprising a potential modulator having a molecular weight of less than 5 kDa.

41. A method for incorporating a potential modulator in a crystal of a polypeptide, comprising placing a crystal of peptidyl-tRNA hydrolase from *P. aeruginosa* having unit cell dimensions $a = 63.702$ Å, $b = 63.702$ Å, $c = 154.844$ Å, $\alpha = \beta = 90°$, $\gamma = 120°$, with an hexagonal space group $P6_122$ in a solution comprising the potential modulator.

42. A computer readable storage medium comprising digitally encoded structural data, wherein the data comprises structural coordinates as listed in FIGURE 9 for the backbone atoms of at least about six amino acid residues from a druggable region of peptidyl-tRNA hydrolase from *P. aeruginosa*.

43. A scalable three-dimensional configuration of points, at least a portion of the points derived from some or all of the structure coordinates as listed in FIGURE 9 for a plurality of amino acid residues from a druggable region of peptidyl-tRNA hydrolase from *P. aeruginosa*.

44. The scalable three-dimensional configuration of points of claim 43, wherein the structure coordinates as listed in FIGURE 9 for the backbone atoms of at least about five amino acid residues from a druggable region of peptidyl-tRNA hydrolase from *P. aeruginosa* are used to derive part or all of the portion of points.

45. The scalable three-dimensional configuration of points of claim 43, wherein the structure coordinates as listed in FIGURE 9 for the backbone and optionally the side chain atoms of at least about ten amino acid residues from a druggable region of peptidyl-tRNA hydrolase from *P. aeruginosa* are used to derive part or all of the portion of points.

46. The scalable three-dimensional configuration of points of claim 43, wherein the structure coordinates as listed in FIGURE 9 for the backbone atoms of at least about fifteen amino acid residues from a druggable region of peptidyl-tRNA hydrolase from *P. aeruginosa* are used to derive part or all of the portion of points.

47. The scalable three-dimensional configuration of points of claim 43, wherein substantially all of the points are derived from structure coordinates as listed in FIGURE 9.

48. The scalable three-dimensional configuration of points of claim 43, wherein the structure coordinates as listed in FIGURE 9 for the atoms of the amino acid residues from any of the above-described druggable regions of peptidyl-tRNA hydrolase from *P. aeruginosa* are used to derive part or all of the portion of points:

49. A scalable three-dimensional configuration of points, comprising points having a root mean square deviation of less than about 1.0 Å from the three dimensional coordinates as listed in FIGURE 9 for the backbone atoms of at least five amino acid residues, wherein the five amino acid residues are from a druggable region of peptidyl-tRNA hydrolase from *P. aeruginosa*.

50. The scalable three-dimensional configuration of points of claim 49, wherein any point-to-point distance, calculated from the three dimensional coordinates as listed in FIGURE 9, between one of the backbone atoms for one of the five amino acid residues and another backbone atom of a different one of the five amino acid residues is not more than about 10 Å.

51. A scalable three-dimensional configuration of points comprising points having a root mean square deviation of less than about 1.0 Å from the three dimensional coordinates as listed in FIGURE 9 for the atoms of the amino acid residues from any of the above-described druggable regions of peptidyl-tRNA hydrolase from *P. aeruginosa*:

52. A computer readable storage medium comprising digitally encoded structural data, wherein the data comprise the identity and three-dimensional coordinates as listed in FIGURE 9 for the atoms of the amino acid residues from any of the above-described druggable regions of peptidyl-tRNA hydrolase from *P. aeruginosa*:

53. A scalable three-dimensional configuration of points, wherein the points have a root mean square deviation of less than about 1.0 Å from the three dimensional coordinates as listed in FIGURE 9 for the atoms of the amino acid residues from any of the above-described druggable regions of peptidyl-tRNA hydrolase from *P. aeruginosa,* wherein up to one amino acid residue in each of the regions may have a conservative substitution thereof.

54. A scalable three-dimensional configuration of points derived from a druggable region of a polypeptide, wherein the points have a root mean square deviation of less than about 1.0 Å from the three dimensional coordinates as listed in FIGURE 9 for the backbone atoms of at least ten amino acid residues that participate in the intersubunit contacts of peptidyl-tRNA hydrolase from *P. aeruginosa*.

55. A computer-assisted method for identifying an inhibitor of the activity of peptidyl-tRNA hydrolase from *P. aeruginosa*, comprising:

(a) supplying a computer modeling application with a set of structure coordinates as listed in FIGURE 9 for the atoms of the amino acid residues from any of the above-described druggable regions of peptidyl-tRNA hydrolase from *P. aeruginosa* so as to define part or all of a molecule or complex;

(b) supplying the computer modeling application with a set of structure coordinates of a chemical entity; and

(c) determining whether the chemical entity is expected to bind to or interfere with the molecule or complex.

56. The method of claim 55, wherein determining whether the chemical entity is expected to bind to or interfere with the molecule or complex comprises performing a fitting operation between the chemical entity and a druggable region of the molecule or complex, followed by computationally analyzing the results of the fitting operation to quantify the association between the chemical entity and the druggable region.

57. The method of claim 55, further comprising screening a library of chemical entities.

58. A computer-assisted method for designing an inhibitor of peptidyl-tRNA hydrolase activity comprising:

(a) supplying a computer modeling application with a set of structure coordinates having a root mean square deviation of less than about 1.0 Å from the structure coordinates as listed in FIGURE 9 for the atoms of the amino acid residues from any of the above-described druggable regions of peptidyl-tRNA hydrolase from *P. aeruginosa* so as to define part or all of a molecule or complex;

(b) supplying the computer modeling application with a set of structure coordinates for a chemical entity;

(c) evaluating the potential binding interactions between the chemical entity and the molecule or complex;

(d) structurally modifying the chemical entity to yield a set of structure coordinates for a modified chemical entity; and

(e) determining whether the modified chemical entity is an inhibitor expected to bind to or interfere with the molecule or complex, wherein binding to or interfering with the molecule or molecular complex is indicative of potential inhibition of peptidyl-tRNA hydrolase activity.

59. The method of claim 58, wherein determining whether the modified chemical entity is an inhibitor expected to bind to or interfere with the molecule or complex comprises performing a fitting operation between the chemical entity and the molecule or complex, followed by computationally analyzing the results of the fitting operation to evaluate the association between the chemical entity and the molecule or complex.

60. The method of claim 58, wherein the set of structure coordinates for the chemical entity is obtained from a chemical library.

61. A computer-assisted method for designing an inhibitor of peptidyl-tRNA hydrolase activity *de novo* comprising:

(a) supplying a computer modeling application with a set of three-dimensional coordinates derived from the structure coordinates as listed in FIGURE 9 for the atoms of the amino acid residues from any of the above-described druggable regions of peptidyl-tRNA hydrolase from *P. aeruginosa* so as to define part or all of a molecule or complex;

(b) computationally building a chemical entity represented by a set of structure coordinates; and

(c) determining whether the chemical entity is an inhibitor expected to bind to or interfere with the molecule or complex, wherein binding to or interfering with the molecule or complex is indicative of potential inhibition of peptidyl-tRNA hydrolase activity.

62. The method of claim 61, wherein determining whether the chemical entity is an inhibitor expected to bind to or interfere with the molecule or complex comprises performing a fitting operation between the chemical entity and a druggable region of the molecule or complex, followed by computationally analyzing the results of the fitting operation to quantify the association between the chemical entity and the druggable region.

63. The method of any of claims 55, 58 or 61, further comprising supplying or synthesizing the potential inhibitor, then assaying the potential inhibitor to determine whether it inhibits peptidyl-tRNA hydrolase activity.

64. A method for identifying a potential modulator for the prevention or treatment of a *P. aeruginosa* related disease or disorder, the method comprising:

(a) providing the three dimensional structure of a crystallized polypeptide comprising: (1) an amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; (2) an amino acid sequence having at least about 95% identity with the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; or (3) an amino acid sequence encoded by a polynucleotide that hybridizes under stringent conditions to the complementary strand of a polynucleotide having SEQ ID NO: 1 or SEQ ID NO: 3 and has at least one biological activity of peptidyl-tRNA hydrolase from *P. aeruginosa*;

(b) obtaining a potential modulator for the prevention or treatment of *P. aeruginosa* related disease or disorder based on the three dimensional structure of the crystallized polypeptide;

(c) contacting the potential modulator with a second polypeptide comprising: (i) an amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; (ii) an amino acid sequence having at least about 95% identity with the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; or (iii) an amino acid sequence encoded by a polynucleotide that hybridizes under stringent conditions to the complementary strand of a polynucleotide having SEQ ID NO: 1 or SEQ ID NO: 3 and has at least one biological activity of peptidyl-

tRNA hydrolase from *P. aeruginosa*; which second polypeptide may optionally be the same as the crystallized polypeptide; and

(d) assaying the activity of the second polypeptide, wherein a change in the activity of the second polypeptide indicates that the compound may be useful for prevention or treatment of a *P. aeruginosa* related disease or disorder.

65. A method for designing a candidate modulator for screening for inhibitors of a polypeptide, the method comprising:

(a) providing the three dimensional structure of a druggable region of a polypeptide comprising (1) an amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; (2) an amino acid sequence having at least about 95% identity with the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; or (3) an amino acid sequence encoded by a polynucleotide that hybridizes under stringent conditions to the complementary strand of a polynucleotide having SEQ ID NO: 1 or SEQ ID NO: 3 and has at least one biological activity of peptidyl-tRNA hydrolase from *P. aeruginosa*; and

(b) designing a candidate modulator based on the three dimensional structure of the druggable region of the polypeptide.

66. A method for identifying a potential modulator of a polypeptide from a database, the method comprising:

(a) providing the three-dimensional coordinates for a plurality of the amino acids of a polypeptide comprising (1) an amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; (2) an amino acid sequence having at least about 95% identity with the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; or (3) an amino acid sequence encoded by a polynucleotide that hybridizes under stringent conditions to the complementary strand of a polynucleotide having SEQ ID NO: 1 or SEQ ID NO: 3 and has at least one biological activity of peptidyl-tRNA hydrolase from *P. aeruginosa*;

(b) identifying a druggable region of the polypeptide; and

(c) selecting from a database at least one potential modulator comprising three dimensional coordinates which indicate that the modulator may bind or interfere with the druggable region.

67. The method of claim 66, wherein the modulator is a small molecule.

68. A method for preparing a potential modulator of a druggable region contained in a polypeptide, the method comprising:

(a) using the atomic coordinates for the backbone atoms of at least about six amino acid residues from a polypeptide of SEQ ID NO: 4, with a ± a root mean square deviation from the backbone atoms of the amino acid residues of not more than 1.0 Å, to generate one or more three-dimensional structures of a molecule comprising a druggable region from the polypeptide;

(b) employing one or more of the three dimensional structures of the molecule to design or select a potential modulator of the druggable region; and

(c) synthesizing or obtaining the modulator.

69. An apparatus for determining whether a compound is a potential modulator of a polypeptide, the apparatus comprising:

(a) a memory that comprises:

(i) the three dimensional coordinates and identities of at least about fifteen atoms from a druggable region of a polypeptide comprising (1) an amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; (2) an amino acid sequence having at least about 95% identity with the amino acid sequence set forth in SEQ ID NO: 2 or SEQ ID NO: 4; or (3) an amino acid sequence encoded by a polynucleotide that hybridizes under stringent conditions to the complementary strand of a polynucleotide having SEQ ID NO: 1 or SEQ ID NO: 3 and has at least one biological activity of peptidyl-tRNA hydrolase from *P. aeruginosa*;

(ii) executable instructions; and

(b) a processor that is capable of executing instructions to:

(i) receive three-dimensional structural information for a candidate modulator;

(ii) determine if the three-dimensional structure of the candidate modulator is complementary to the three dimensional coordinates of the atoms from the druggable region; and

(iii) output the results of the determination.

70. A method for making an inhibitor of peptidyl-tRNA hydrolase activity, the method comprising chemically or enzymatically synthesizing a chemical entity to yield an

inhibitor of peptidyl-tRNA hydrolase activity, the chemical entity having been identified during a computer-assisted process comprising supplying a computer modeling application with a set of structure coordinates of a molecule or complex, the molecule or complex comprising at least a portion of at least one druggable region from peptidyl-tRNA hydrolase from *P. aeruginosa*; supplying the computer modeling application with a set of structure coordinates of a chemical entity; and determining whether the chemical entity is expected to bind or to interfere with the molecule or complex at a druggable region, wherein binding to or interfering with the molecule or complex is indicative of potential inhibition of peptidyl-tRNA hydrolase activity.

71. A computer readable storage medium comprising digitally encoded data, wherein the data comprises structural coordinates for a druggable region that is structurally homologous to the structure coordinates as listed in FIGURE 9 for a druggable region of peptidyl-tRNA hydrolase from *P. aeruginosa*.

72. A computer readable storage medium comprising digitally encoded structural data, wherein the data comprise a majority of the three-dimensional structure coordinates as listed in FIGURE 9.

73. The computer readable storage medium of claim 72, further comprising the identity of the atoms for the majority of the three-dimensional structure coordinates as listed in FIGURE 9.

The computer readable storage medium of claim 72, wherein the data comprise substantially all of the three-dimensional structure coordinates as listed in FIGURE 9.

FIGURE 1

SEQ ID NO: 1

GTGACTGCCGTACAACTGATCGTTGGCCTGGGAAACCCGGGTCCTGAAT
ACGACCAGACCCGGCATAACGCGGGGGCCCTTTTCGTTGAGCGCCTGGCGCAT
GCCCAGGGCGTCAGCCTCGTGGCTGACCGCAAGTATTTCGGCCTGGTCGGCAA
GTTCAGCCACCAGGGCAAGGACGTTCGTCTGTTGATCCCGACCACCTACATGAA
CCGCAGCGGCCAGTCCGTGGCGGCGCTGGCGGGATTCTTCCGGATCGCCCCGG
ACGCCATCCTGGTGGCCCACGACGAACTCGACATGCCCCCTGGCGTCGCCAAG
CTCAAGACCGGCGGCGGACACGGCGGGCACAACGGGTTGCGCGACATCATCGC
CCAGCTCGGCAATCAGAATTCTTTCCATCGCCTGCGGCTTGGCATCGGCCATCC
GGGGCACAGCAGCCTGGTTTCCGGTTACGTGCTCGGCCGCGCCCCGCGCAGCG
AGCAGGAACTGCTCGACACCAGCATCGACTTCGCCCTCGGCGTGCTGCCGGAA
ATGCTCGCCGGGGACTGGACCCGGGCGATGCAGAAGCTGCACAGCCAGAAGGC
CTGA

FIGURE 2

SEQ ID NO: 2

VTAVQLIVGLGNPGPEYDQTRHNAGALFVERLAHAQGVSLVADRKYFGLV
GKFSHQGKDVRLLIPTTYMNRSGQSVAALAGFFRIAPDAILVAHDELDMPPGVAKL
KTGGGHGGHNGLRDIIAQLGNQNSFHRLRLGIGHPGHSSLVSGYVLGRAPRSEQEL
LDTSIDFALGVLPEMLAGDWTRAMQKLHSQKA

FIGURE 3

SEQ ID NO: 3

GTGACTGCCGTACAACTGATCGTTGGCCTGGGAAACCCGGGTCCTGAAT
ACGACCAGACCCGGCATAACGCGGGGGCCCTTTTCGTTGAGCGCCTGGCGCAT
GCCCAGGGCGTCAGCCTCGTGGCTGACCGCAAGTATTTCGGCCTGGTCGGCAA
GTTCAGCCACCAGGGCAAGGACGTTCGTCTGTTGATCCCGACCACCTACATGAA
CCGCAGCGGCCAGTCCGTGGCGGCGCTGGCGGGATTCTTCCGGATCGCCCCGG
ACGCCATCCTGGTGGCCCACGACGAACTCGACATGCCCCCTGGCGTCGCCAAG
CTCAAGACCGGCGGCGGACACGGCGGGCACAACGGGTTGCGCGACATCATCGC
CCAGCTCGGCAATCAGAATTCTTTCCATCGCCTGCGGCTTGGCATCGGCCATCC
GGGGCACAGCAGCCTGGTTTCCGGTTACGTGCTCGGCCGCGCCCCGCGCAGCG
AGCAGGAACTGCTCGACACCAGCATCGACTTCGCCCTCGGCGTGCTGCCGGAA
ATGCTCGCCGGGGACTGGACCCGGGCGATGCAGAAGCTGCACAGCCAGAAGGC
CTGA

FIGURE 4

SEQ ID NO: 4

VTAVQLIVGLGNPGPEYDQTRHNAGALFVERLAHAQGVSLVADRKYFGLV
GKFSHQGKDVRLLIPTTYMNRSGQSVAALAGFFRIAPDAILVAHDELDMPPGVAKL
KTGGGHGGHNGLRDIIAQLGNQNSFHRLRLGIGHPGHSSLVSGYVLGRAPRSEQEL
LDTSIDFALGVLPEMLAGDWTRAMQKLHSQKA

FIGURE 5

SEQ ID NO: 5

Forward PCR Primer

GCGGCGGCCCATATGACTGCCGTACAACTGATC

SEQ ID NO: 6

Reverse PCR Primer

GCGCGGATCCGGCCTTCTGGCTGTGCAG

FIGURE 6

TABLE 1: Amino Acid and Nucleic Acid Properties

Melting temperature (°C) of SEQ ID NO: 5 (forward PCR primer)	62
Restriction enzyme for SEQ ID NO: 5 (forward PCR primer)	NdeI
Melting temperature (°C) of SEQ ID NO: 6 (reverse PCR primer)	60
Restriction enzyme for SEQ ID NO: 6 (reverse PCR primer)	BamHI
Number of nucleic acid residues in SEQ ID NO: 1	585
Number of amino acid residues in SEQ ID NO: 2	194
Number of different nucleic acid residues between SEQ ID NO: 1 and SEQ ID NO: 3	0
Number of different amino acid residues between SEQ ID NO: 2 and SEQ ID NO: 4	0
Calculated molecular weight of SEQ ID NO: 2 polypeptide (kDa)	20.804
Calculated pI of SEQ ID NO: 2 polypeptide	10
Solubility of SEQ ID NO: 4 polypeptide, determined as described in EXAMPLE 2 (with the His tag at the N-terminus)	Approximately one third
Solubility of SEQ ID NO: 4 polypeptide, determined as described in EXAMPLE 2 (with the His tag at the C-terminus)	No detectable expression
Amount of purified polypeptide having SEQ ID NO: 4, prepared and purified as described in EXAMPLE 8 (mg/L of culture)	3.9
Amount of purified polypeptide having SEQ ID NO: 4 soluble in buffer, as described in EXAMPLE 8 (mg/ml of buffer)	11

FIGURE 7

TABLE 2: Bioinformatic Analyses

Protein annotation and gene designation, if any	peptidyl-tRNA hydrolase, *pth*
COG Category	translation, ribosomal structure and biogenesis
COG ID Number	COG0193
Is SEQ ID NO: 2 classified as an essential gene?	yes
Most closely related protein from PDB	peptidyl-tRNA hydrolase (2pth)
Source organism for closest PDB protein	*Escherichia coli*
e-value for closest PDB Protein	58
% Identity between SEQ ID NO: 2 and the closest protein from PDB	72
% Positives between SEQ ID NO: 2 and the closest protein from PDB	4.00E-60
Number of Protein Hits in the VGDB	13
Number of Microorganisms having VGDB Hits	13
Microorganisms having VGDB Hits[1]	[paer][hinf][bsub][nmen][ecoli][spne][saur] [efae][mgen][ctra][bbur][rpxx][hpyl]
First predicted epitopic region of SEQ ID NO: 2: rank score, amino acid residue numbers; amino acid sequence	1.188, 133->155, SEQ ID NO: 10: RLRLGIGHPGHSSLVSGYVLGRA
Second predicted epitopic region of SEQ ID NO: 2: rank score, amino acid residue numbers; amino acid sequence	1.187, 4->11, SEQ ID NO: 11: VQLIVGLG
Third predicted epitopic region of SEQ ID NO: 2: rank score, amino acid residue numbers; amino acid sequence	1.140, 24->44, SEQ ID NO: 12: AGALFVERLAHAQGVSLVADR

[1]Organisms are abbreviated as follows: ecoli = *Eschericia coli*; hpyl = *Helicobacter pylori*; paer = *Pseudomonas aeruginosa*; ctra = *Chlaydia trachomatis*; hinf = *Haemophilus influenzae*; nmen = *Neisseria meningitidis*; rpxx = *Rickettsia prowazekii*; bbur = *Borrelia burgdorferi*; bsub = *Bacillus subtilis*; staph = *Staphylococcus aureus*; spne = *Streptococcus pneumoniae*; mgen = *Mycoplasma genitalium*; efae = *Enterococcus faecalis*.

FIGURE 8

TABLE 3: X-ray Structure Data

Data collection	
Wavelength	1.54
Space group	P6(1)22
Unit cell dimensions (Å)	$a = b = 63.702$, $c = 154.844$; $= 90$ $= 90$, $= 120$
Resolution range (Å)	30-1.90
Completeness (%)[a,b]	99.7 (99.8)
$I/\ I$	9.4 (2.93)
R_{merge}	6.3 (27.8)
Number of reflections	
Total	152832
Unique	15497
Structure Refinement	
R_{cryst}[d]	18.8
R_{free}	22.8
Number of atoms	
protein	1453
solvent	215
Average B-factors	
protein (Å^2)	14.4
solvent (Å^2)	27
R.M.S. Deviations from ideal	
Bond (Å)	0.0044
Angle (°)	1.14

[a]Number in parentheses is the statistic for highest resolution shell.
[b]$I \geq \sigma I$
[c]$R_{sym} = \Sigma_h(\Sigma_j | I_{j,h} - < I_h > | / \Sigma I_{j,h})$, where h = set of Miller indices and j = set of observations of reflection h.
[d]$R_{cryst} = \Sigma_{hkl} | F_o - F_d | / \Sigma_{hkl} | F_o |$

FIGURE 9

```
ATOM      1  CB  THR   1A    -47.075  90.346 -66.456  1.00 13.53 C
ATOM      2  OG1 THR   1A    -47.543  90.425 -67.808  1.00 16.07 O
ATOM      3  CG2 THR   1A    -47.599  91.530 -65.664  1.00 17.49 C
ATOM      4  C   THR   1A    -45.020  89.060 -67.094  1.00  6.29 C
ATOM      5  O   THR   1A    -44.698  89.061 -68.279  1.00  5.51 O
ATOM      6  N   THR   1A    -44.960  91.498 -67.080  1.00  9.53 N
ATOM      7  CA  THR   1A    -45.532  90.330 -66.426  1.00 10.29 C
ATOM      8  N   ILE   2A    -44.933  87.985 -66.320  1.00  3.48 N
ATOM      9  CA  ILE   2A    -44.465  86.707 -66.840  1.00  3.08 C
ATOM     10  CB  ILE   2A    -43.996  85.789 -65.691  1.00  2.74 C
ATOM     11  CG2 ILE   2A    -43.755  84.373 -66.206  1.00  4.38 C
ATOM     12  CG1 ILE   2A    -42.742  86.381 -65.052  1.00  3.38 C
ATOM     13  CD1 ILE   2A    -41.630  86.632 -66.037  1.00 10.37 C
ATOM     14  C   ILE   2A    -45.562  86.003 -67.629  1.00  3.20 C
ATOM     15  O   ILE   2A    -46.696  85.879 -67.156  1.00  3.27 O
ATOM     16  N   LYS   3A    -45.212  85.544 -68.827  1.00  3.01 N
ATOM     17  CA  LYS   3A    -46.149  84.845 -69.697  1.00  6.53 C
ATOM     18  CB  LYS   3A    -46.243  85.564 -71.047  1.00 11.47 C
ATOM     19  CG  LYS   3A    -46.781  86.990 -70.947  1.00 18.25 C
ATOM     20  CD  LYS   3A    -46.818  87.676 -72.309  1.00 23.70 C
ATOM     21  CE  LYS   3A    -47.752  86.949 -73.262  1.00 28.19 C
ATOM     22  NZ  LYS   3A    -49.130  86.845 -72.690  1.00 29.16 N
ATOM     23  C   LYS   3A    -45.758  83.381 -69.914  1.00  3.71 C
ATOM     24  O   LYS   3A    -46.511  82.614 -70.510  1.00  1.00 O
ATOM     25  N   LEU   4A    -44.582  82.997 -69.423  1.00  2.05 N
ATOM     26  CA  LEU   4A    -44.104  81.627 -69.565  1.00  2.22 C
ATOM     27  CB  LEU   4A    -43.363  81.459 -70.899  1.00  1.89 C
ATOM     28  CG  LEU   4A    -42.761  80.079 -71.176  1.00  3.80 C
ATOM     29  CD1 LEU   4A    -43.881  79.048 -71.239  1.00  3.64 C
ATOM     30  CD2 LEU   4A    -41.968  80.096 -72.485  1.00  1.51 C
ATOM     31  C   LEU   4A    -43.188  81.198 -68.416  1.00  1.19 C
ATOM     32  O   LEU   4A    -42.206  81.871 -68.092  1.00  1.00 O
ATOM     33  N   ILE   5A    -43.531  80.075 -67.794  1.00  1.87 N
ATOM     34  CA  ILE   5A    -42.732  79.525 -66.712  1.00  1.00 C
ATOM     35  CB  ILE   5A    -43.525  79.426 -65.392  1.00  1.00 C
ATOM     36  CG2 ILE   5A    -42.635  78.793 -64.314  1.00  1.51 C
ATOM     37  CG1 ILE   5A    -43.967  80.823 -64.944  1.00  2.97 C
ATOM     38  CD1 ILE   5A    -44.883  80.829 -63.726  1.00  3.09 C
ATOM     39  C   ILE   5A    -42.287  78.134 -67.160  1.00  1.28 C
ATOM     40  O   ILE   5A    -43.110  77.265 -67.452  1.00  1.00 O
ATOM     41  N   VAL   6A    -40.974  77.946 -67.231  1.00  1.00 N
ATOM     42  CA  VAL   6A    -40.392  76.688 -67.673  1.00  1.32 C
ATOM     43  CB  VAL   6A    -39.310  76.941 -68.752  1.00  2.99 C
ATOM     44  CG1 VAL   6A    -38.905  75.628 -69.409  1.00  2.74 C
ATOM     45  CG2 VAL   6A    -39.819  77.924 -69.788  1.00  3.16 C
ATOM     46  C   VAL   6A    -39.753  75.938 -66.502  1.00  1.53 C
ATOM     47  O   VAL   6A    -39.081  76.533 -65.663  1.00  2.65 O
ATOM     48  N   GLY   7A    -39.977  74.630 -66.457  1.00  1.74 N
ATOM     49  CA  GLY   7A    -39.423  73.804 -65.400  1.00  1.00 C
ATOM     50  C   GLY   7A    -38.565  72.741 -66.044  1.00  2.89 C
ATOM     51  O   GLY   7A    -39.080  71.810 -66.670  1.00  1.00 O
ATOM     52  N   LEU   8A    -37.253  72.893 -65.887  1.00  1.00 N
ATOM     53  CA  LEU   8A    -36.264  71.998 -66.479  1.00  1.00 C
ATOM     54  CB  LEU   8A    -34.899  72.689 -66.490  1.00  1.52 C
ATOM     55  CG  LEU   8A    -34.816  74.032 -67.223  1.00  3.81 C
ATOM     56  CD1 LEU   8A    -33.363  74.505 -67.261  1.00  2.62 C
ATOM     57  CD2 LEU   8A    -35.361  73.882 -68.639  1.00  4.14 C
ATOM     58  C   LEU   8A    -36.125  70.615 -65.850  1.00  1.31 C
```

FIGURE 9 -1

ATOM	59	O	LEU	8A	-36.184	70.455	-64.628	1.00	1.71	O
ATOM	60	N	ALA	9A	-35.920	69.618	-66.703	1.00	1.89	N
ATOM	61	CA	ALA	9A	-35.751	68.245	-66.242	1.00	3.38	C
ATOM	62	CB	ALA	9A	-36.987	67.792	-65.507	1.00	6.64	C
ATOM	63	C	ALA	9A	-35.491	67.317	-67.412	1.00	5.53	C
ATOM	64	O	ALA	9A	-35.545	67.733	-68.569	1.00	7.86	O
ATOM	65	N	ASN	10A	-35.187	66.064	-67.094	1.00	4.29	N
ATOM	66	CA	ASN	10A	-34.960	65.042	-68.104	1.00	4.39	C
ATOM	67	CB	ASN	10A	-33.736	64.185	-67.763	1.00	3.09	C
ATOM	68	CG	ASN	10A	-32.426	64.850	-68.148	1.00	3.78	C
ATOM	69	OD1	ASN	10A	-32.187	65.144	-69.317	1.00	5.67	O
ATOM	70	ND2	ASN	10A	-31.571	65.074	-67.167	1.00	6.15	N
ATOM	71	C	ASN	10A	-36.216	64.174	-68.087	1.00	5.86	C
ATOM	72	O	ASN	10A	-36.851	64.009	-67.048	1.00	4.43	O
ATOM	73	N	PRO	11A	-36.606	63.631	-69.245	1.00	6.76	N
ATOM	74	CD	PRO	11A	-36.140	63.922	-70.613	1.00	7.66	C
ATOM	75	CA	PRO	11A	-37.804	62.795	-69.265	1.00	7.50	C
ATOM	76	CB	PRO	11A	-38.319	62.992	-70.678	1.00	8.92	C
ATOM	77	CG	PRO	11A	-37.051	63.050	-71.465	1.00	7.92	C
ATOM	78	C	PRO	11A	-37.446	61.342	-68.976	1.00	6.42	C
ATOM	79	O	PRO	11A	-36.296	60.947	-69.134	1.00	7.32	O
ATOM	80	N	GLY	12A	-38.432	60.566	-68.548	1.00	7.68	N
ATOM	81	CA	GLY	12A	-38.196	59.162	-68.268	1.00	7.99	C
ATOM	82	C	GLY	12A	-38.285	58.776	-66.806	1.00	7.65	C
ATOM	83	O	GLY	12A	-37.994	59.578	-65.915	1.00	6.95	O
ATOM	84	N	PRO	13A	-38.679	57.525	-66.531	1.00	7.42	N
ATOM	85	CD	PRO	13A	-39.200	56.549	-67.505	1.00	8.54	C
ATOM	86	CA	PRO	13A	-38.809	57.017	-65.166	1.00	7.67	C
ATOM	87	CB	PRO	13A	-39.275	55.579	-65.382	1.00	6.90	C
ATOM	88	CG	PRO	13A	-40.095	55.685	-66.646	1.00	4.45	C
ATOM	89	C	PRO	13A	-37.514	57.095	-64.362	1.00	8.84	C
ATOM	90	O	PRO	13A	-37.527	57.459	-63.187	1.00	9.19	O
ATOM	91	N	GLU	14A	-36.397	56.771	-65.003	1.00	10.22	N
ATOM	92	CA	GLU	14A	-35.109	56.782	-64.315	1.00	11.61	C
ATOM	93	CB	GLU	14A	-34.081	55.988	-65.123	1.00	15.70	C
ATOM	94	CG	GLU	14A	-34.530	54.564	-65.468	1.00	25.30	C
ATOM	95	CD	GLU	14A	-34.849	53.715	-64.242	1.00	30.12	C
ATOM	96	OE1	GLU	14A	-35.781	54.073	-63.488	1.00	33.10	O
ATOM	97	OE2	GLU	14A	-34.169	52.685	-64.035	1.00	32.34	O
ATOM	98	C	GLU	14A	-34.555	58.169	-63.993	1.00	8.01	C
ATOM	99	O	GLU	14A	-33.566	58.291	-63.273	1.00	7.91	O
ATOM	100	N	TYR	15A	-35.191	59.213	-64.511	1.00	6.95	N
ATOM	101	CA	TYR	15A	-34.731	60.572	-64.241	1.00	6.80	C
ATOM	102	CB	TYR	15A	-34.606	61.350	-65.554	1.00	6.12	C
ATOM	103	CG	TYR	15A	-33.490	60.872	-66.455	1.00	5.11	C
ATOM	104	CD1	TYR	15A	-32.166	61.199	-66.184	1.00	6.58	C
ATOM	105	CE1	TYR	15A	-31.134	60.753	-67.004	1.00	8.12	C
ATOM	106	CD2	TYR	15A	-33.760	60.084	-67.577	1.00	6.92	C
ATOM	107	CE2	TYR	15A	-32.733	59.632	-68.406	1.00	7.63	C
ATOM	108	CZ	TYR	15A	-31.424	59.969	-68.112	1.00	8.00	C
ATOM	109	OH	TYR	15A	-30.400	59.517	-68.914	1.00	9.35	O
ATOM	110	C	TYR	15A	-35.702	61.303	-63.320	1.00	8.78	C
ATOM	111	O	TYR	15A	-35.325	62.231	-62.614	1.00	9.47	O
ATOM	112	N	ASP	16A	-36.952	60.859	-63.331	1.00	9.97	N
ATOM	113	CA	ASP	16A	-38.017	61.475	-62.550	1.00	11.63	C
ATOM	114	CB	ASP	16A	-39.061	60.424	-62.172	1.00	11.87	C
ATOM	115	CG	ASP	16A	-40.354	61.047	-61.695	1.00	7.19	C
ATOM	116	OD1	ASP	16A	-41.223	60.313	-61.197	1.00	11.02	O
ATOM	117	OD2	ASP	16A	-40.504	62.277	-61.827	1.00	11.93	O
ATOM	118	C	ASP	16A	-37.630	62.279	-61.303	1.00	12.31	C
ATOM	119	O	ASP	16A	-37.500	63.500	-61.368	1.00	15.69	O

FIGURE 9 - 2

```
ATOM    120  N   GLN   17A   -37.438  61.600 -60.176  1.00 12.79 N
ATOM    121  CA  GLN   17A   -37.115  62.278 -58.921  1.00 10.92 C
ATOM    122  CB  GLN   17A   -37.626  61.441 -57.743  1.00 14.22 C
ATOM    123  CG  GLN   17A   -39.093  61.024 -57.853  1.00 17.50 C
ATOM    124  CD  GLN   17A   -40.049  62.206 -57.882  1.00 21.87 C
ATOM    125  OE1 GLN   17A   -40.107  63.001 -56.941  1.00 25.25 O
ATOM    126  NE2 GLN   17A   -40.807  62.326 -58.965  1.00 21.57 N
ATOM    127  C   GLN   17A   -35.642  62.632 -58.687  1.00  9.05 C
ATOM    128  O   GLN   17A   -35.199  62.719 -57.541  1.00  7.82 O
ATOM    129  N   THR   18A   -34.886  62.849 -59.757  1.00  5.77 N
ATOM    130  CA  THR   18A   -33.474  63.194 -59.598  1.00  3.88 C
ATOM    131  CB  THR   18A   -32.654  62.797 -60.840  1.00  4.24 C
ATOM    132  OG1 THR   18A   -33.177  63.461 -61.994  1.00  3.19 O
ATOM    133  CG2 THR   18A   -32.716  61.285 -61.060  1.00  3.10 C
ATOM    134  C   THR   18A   -33.320  64.690 -59.331  1.00  3.07 C
ATOM    135  O   THR   18A   -34.208  65.486 -59.644  1.00  1.17 O
ATOM    136  N   ARG   19A   -32.194  65.070 -58.737  1.00  1.34 N
ATOM    137  CA  ARG   19A   -31.953  66.462 -58.407  1.00  2.22 C
ATOM    138  CB  ARG   19A   -30.561  66.621 -57.778  1.00  1.00 C
ATOM    139  CG  ARG   19A   -30.324  65.848 -56.486  1.00  1.00 C
ATOM    140  CD  ARG   19A   -28.941  66.210 -55.925  1.00  2.02 C
ATOM    141  NE  ARG   19A   -28.664  65.617 -54.620  1.00  3.66 N
ATOM    142  CZ  ARG   19A   -28.011  64.475 -54.427  1.00  5.49 C
ATOM    143  NH1 ARG   19A   -27.554  63.777 -55.460  1.00  8.27 N
ATOM    144  NH2 ARG   19A   -27.796  64.036 -53.194  1.00  4.56 N
ATOM    145  C   ARG   19A   -32.056  67.394 -59.613  1.00  1.00 C
ATOM    146  O   ARG   19A   -32.522  68.527 -59.488  1.00  3.32 O
ATOM    147  N   HIS   20A   -31.612  66.913 -60.771  1.00  1.00 N
ATOM    148  CA  HIS   20A   -31.604  67.711 -61.993  1.00  1.00 C
ATOM    149  CB  HIS   20A   -30.816  66.969 -63.078  1.00  2.69 C
ATOM    150  CG  HIS   20A   -30.099  67.872 -64.035  1.00  4.77 C
ATOM    151  CD2 HIS   20A   -29.871  67.763 -65.364  1.00  5.27 C
ATOM    152  ND1 HIS   20A   -29.462  69.027 -63.632  1.00  4.39 N
ATOM    153  CE1 HIS   20A   -28.872  69.587 -64.672  1.00  4.22 C
ATOM    154  NE2 HIS   20A   -29.104  68.840 -65.737  1.00  3.23 N
ATOM    155  C   HIS   20A   -33.003  68.073 -62.488  1.00  1.84 C
ATOM    156  O   HIS   20A   -33.145  68.889 -63.397  1.00  1.00 O
ATOM    157  N   ASN   21A   -34.029  67.471 -61.880  1.00  1.00 N
ATOM    158  CA  ASN   21A   -35.419  67.739 -62.242  1.00  1.72 C
ATOM    159  CB  ASN   21A   -36.221  66.431 -62.289  1.00  1.00 C
ATOM    160  CG  ASN   21A   -36.039  65.691 -63.589  1.00  1.00 C
ATOM    161  OD1 ASN   21A   -34.979  65.764 -64.211  1.00  1.00 O
ATOM    162  ND2 ASN   21A   -37.074  64.967 -64.009  1.00  2.66 N
ATOM    163  C   ASN   21A   -36.099  68.703 -61.277  1.00  1.68 C
ATOM    164  O   ASN   21A   -37.324  68.777 -61.231  1.00  1.00 O
ATOM    165  N   ALA   22A   -35.308  69.436 -60.498  1.00  1.00 N
ATOM    166  CA  ALA   22A   -35.869  70.382 -59.543  1.00  1.96 C
ATOM    167  CB  ALA   22A   -34.749  71.161 -58.859  1.00  5.70 C
ATOM    168  C   ALA   22A   -36.840  71.347 -60.220  1.00  1.00 C
ATOM    169  O   ALA   22A   -37.944  71.588 -59.719  1.00  2.92 O
ATOM    170  N   GLY   23A   -36.424  71.896 -61.358  1.00  1.00 N
ATOM    171  CA  GLY   23A   -37.255  72.840 -62.086  1.00  1.45 C
ATOM    172  C   GLY   23A   -38.602  72.271 -62.484  1.00  1.00 C
ATOM    173  O   GLY   23A   -39.636  72.912 -62.298  1.00  1.00 O
ATOM    174  N   ALA   24A   -38.597  71.060 -63.033  1.00  1.00 N
ATOM    175  CA  ALA   24A   -39.841  70.428 -63.450  1.00  1.00 C
ATOM    176  CB  ALA   24A   -39.551  69.101 -64.151  1.00  1.68 C
ATOM    177  C   ALA   24A   -40.761  70.194 -62.261  1.00  2.85 C
ATOM    178  O   ALA   24A   -41.986  70.287 -62.390  1.00  1.00 O
ATOM    179  N   LEU   25A   -40.176  69.883 -61.108  1.00  1.82 N
ATOM    180  CA  LEU   25A   -40.964  69.641 -59.906  1.00  2.74 C
```

FIGURE 9 - 3

```
ATOM    181  CB  LEU   25A    -40.080  69.067 -58.787  1.00  2.83 C
ATOM    182  CG  LEU   25A    -40.781  68.747 -57.462  1.00  1.89 C
ATOM    183  CD1 LEU   25A    -41.790  67.626 -57.665  1.00  2.65 C
ATOM    184  CD2 LEU   25A    -39.741  68.332 -56.424  1.00  2.70 C
ATOM    185  C   LEU   25A    -41.622  70.943 -59.459  1.00  3.26 C
ATOM    186  O   LEU   25A    -42.772  70.950 -59.025  1.00  1.00 O
ATOM    187  N   PHE   26A    -40.898  72.052 -59.565  1.00  1.77 N
ATOM    188  CA  PHE   26A    -41.480  73.322 -59.181  1.00  1.86 C
ATOM    189  CB  PHE   26A    -40.483  74.466 -59.356  1.00  1.00 C
ATOM    190  CG  PHE   26A    -41.107  75.816 -59.231  1.00  2.24 C
ATOM    191  CD1 PHE   26A    -41.443  76.325 -57.984  1.00  1.24 C
ATOM    192  CD2 PHE   26A    -41.406  76.561 -60.365  1.00  1.00 C
ATOM    193  CE1 PHE   26A    -42.075  77.562 -57.868  1.00  1.00 C
ATOM    194  CE2 PHE   26A    -42.038  77.796 -60.256  1.00  1.33 C
ATOM    195  CZ  PHE   26A    -42.373  78.293 -59.005  1.00  1.00 C
ATOM    196  C   PHE   26A    -42.707  73.608 -60.036  1.00  1.57 C
ATOM    197  O   PHE   26A    -43.740  74.011 -59.518  1.00  1.52 O
ATOM    198  N   VAL   27A    -42.593  73.402 -61.345  1.00  1.15 N
ATOM    199  CA  VAL   27A    -43.717  73.659 -62.242  1.00  1.00 C
ATOM    200  CB  VAL   27A    -43.254  73.690 -63.726  1.00  1.00 C
ATOM    201  CG1 VAL   27A    -44.455  73.881 -64.655  1.00  5.73 C
ATOM    202  CG2 VAL   27A    -42.273  74.835 -63.931  1.00  2.04 C
ATOM    203  C   VAL   27A    -44.877  72.675 -62.078  1.00  1.68 C
ATOM    204  O   VAL   27A    -46.034  73.036 -62.305  1.00  1.00 O
ATOM    205  N   GLU   28A    -44.578  71.441 -61.681  1.00  1.07 N
ATOM    206  CA  GLU   28A    -45.636  70.457 -61.478  1.00  1.48 C
ATOM    207  CB  GLU   28A    -45.044  69.060 -61.256  1.00  1.00 C
ATOM    208  CG  GLU   28A    -46.070  67.954 -61.396  1.00  7.79 C
ATOM    209  CD  GLU   28A    -45.501  66.581 -61.119  1.00  4.86 C
ATOM    210  OE1 GLU   28A    -44.369  66.295 -61.558  1.00  7.74 O
ATOM    211  OE2 GLU   28A    -46.200  65.786 -60.462  1.00  9.93 O
ATOM    212  C   GLU   28A    -46.466  70.880 -60.260  1.00  2.25 C
ATOM    213  O   GLU   28A    -47.691  70.748 -60.254  1.00  3.18 O
ATOM    214  N   ARG   29A    -45.790  71.391 -59.233  1.00  1.00 N
ATOM    215  CA  ARG   29A    -46.461  71.849 -58.018  1.00  2.69 C
ATOM    216  CB  ARG   29A    -45.440  72.190 -56.929  1.00  6.04 C
ATOM    217  CG  ARG   29A    -44.671  71.006 -56.381  1.00  7.97 C
ATOM    218  CD  ARG   29A    -45.556  70.082 -55.556  1.00 11.91 C
ATOM    219  NE  ARG   29A    -44.765  69.065 -54.869  1.00 13.66 N
ATOM    220  CZ  ARG   29A    -45.251  68.196 -53.986  1.00 13.68 C
ATOM    221  NH1 ARG   29A    -46.539  68.206 -53.672  1.00 15.72 N
ATOM    222  NH2 ARG   29A    -44.439  67.325 -53.405  1.00 15.39 N
ATOM    223  C   ARG   29A    -47.277  73.092 -58.323  1.00  2.80 C
ATOM    224  O   ARG   29A    -48.419  73.222 -57.881  1.00  2.22 O
ATOM    225  N   LEU   30A    -46.671  74.009 -59.069  1.00  1.00 N
ATOM    226  CA  LEU   30A    -47.333  75.245 -59.459  1.00  1.44 C
ATOM    227  CB  LEU   30A    -46.381  76.105 -60.302  1.00  2.02 C
ATOM    228  CG  LEU   30A    -46.840  77.510 -60.730  1.00  1.00 C
ATOM    229  CD1 LEU   30A    -45.684  78.231 -61.413  1.00  6.09 C
ATOM    230  CD2 LEU   30A    -48.021  77.413 -61.673  1.00  5.54 C
ATOM    231  C   LEU   30A    -48.589  74.915 -60.263  1.00  2.84 C
ATOM    232  O   LEU   30A    -49.661  75.483 -60.027  1.00  2.47 O
ATOM    233  N   ALA   31A    -48.455  73.989 -61.208  1.00  1.89 N
ATOM    234  CA  ALA   31A    -49.584  73.604 -62.048  1.00  2.26 C
ATOM    235  CB  ALA   31A    -49.135  72.612 -63.112  1.00  2.51 C
ATOM    236  C   ALA   31A    -50.721  73.008 -61.219  1.00  1.67 C
ATOM    237  O   ALA   31A    -51.887  73.335 -61.427  1.00  1.00 O
ATOM    238  N   HIS   32A    -50.375  72.147 -60.270  1.00  1.00 N
ATOM    239  CA  HIS   32A    -51.381  71.516 -59.424  1.00  3.65 C
ATOM    240  CB  HIS   32A    -50.750  70.417 -58.561  1.00  7.40 C
ATOM    241  CG  HIS   32A    -51.750  69.634 -57.774  1.00 10.18 C
```

FIGURE 9 - 4

```
ATOM    242  CD2 HIS    32A     -51.962  69.541 -56.440  1.00 12.23 C
ATOM    243  ND1 HIS    32A     -52.724  68.864 -58.371  1.00 12.96 N
ATOM    244  CE1 HIS    32A     -53.494  68.330 -57.440  1.00 13.57 C
ATOM    245  NE2 HIS    32A     -53.053  68.726 -56.259  1.00 14.32 N
ATOM    246  C   HIS    32A     -52.066  72.541 -58.524  1.00  5.16 C
ATOM    247  O   HIS    32A     -53.285  72.504 -58.341  1.00  3.04 O
ATOM    248  N   ALA    33A     -51.273  73.454 -57.971  1.00  5.26 N
ATOM    249  CA  ALA    33A     -51.796  74.496 -57.094  1.00  4.69 C
ATOM    250  CB  ALA    33A     -50.637  75.287 -56.473  1.00  2.35 C
ATOM    251  C   ALA    33A     -52.737  75.436 -57.849  1.00  5.73 C
ATOM    252  O   ALA    33A     -53.690  75.963 -57.270  1.00  4.91 O
ATOM    253  N   GLN    34A     -52.482  75.629 -59.141  1.00  6.05 N
ATOM    254  CA  GLN    34A     -53.308  76.514 -59.959  1.00  7.15 C
ATOM    255  CB  GLN    34A     -52.420  77.326 -60.919  1.00  6.75 C
ATOM    256  CG  GLN    34A     -51.403  78.229 -60.235  1.00 10.76 C
ATOM    257  CD  GLN    34A     -52.056  79.284 -59.363  1.00 12.81 C
ATOM    258  OE1 GLN    34A     -52.945  80.010 -59.808  1.00 14.82 O
ATOM    259  NE2 GLN    34A     -51.614  79.376 -58.112  1.00 14.83 N
ATOM    260  C   GLN    34A     -54.390  75.780 -60.759  1.00  6.33 C
ATOM    261  O   GLN    34A     -55.068  76.386 -61.590  1.00  6.53 O
ATOM    262  N   GLY    35A     -54.540  74.481 -60.507  1.00  5.57 N
ATOM    263  CA  GLY    35A     -55.539  73.680 -61.204  1.00  5.81 C
ATOM    264  C   GLY    35A     -55.314  73.592 -62.706  1.00  7.61 C
ATOM    265  O   GLY    35A     -56.259  73.418 -63.485  1.00  7.32 O
ATOM    266  N   VAL    36A     -54.052  73.691 -63.112  1.00  5.36 N
ATOM    267  CA  VAL    36A     -53.685  73.635 -64.521  1.00  3.92 C
ATOM    268  CB  VAL    36A     -52.515  74.598 -64.824  1.00  3.54 C
ATOM    269  CG1 VAL    36A     -52.067  74.436 -66.277  1.00  4.84 C
ATOM    270  CG2 VAL    36A     -52.929  76.033 -64.536  1.00  2.36 C
ATOM    271  C   VAL    36A     -53.259  72.245 -64.980  1.00  3.98 C
ATOM    272  O   VAL    36A     -52.405  71.608 -64.358  1.00  2.89 O
ATOM    273  N   SER    37A     -53.847  71.793 -66.081  1.00  3.52 N
ATOM    274  CA  SER    37A     -53.512  70.494 -66.652  1.00  4.67 C
ATOM    275  CB  SER    37A     -54.707  69.920 -67.424  1.00  5.75 C
ATOM    276  OG  SER    37A     -55.760  69.528 -66.557  1.00  7.63 O
ATOM    277  C   SER    37A     -52.343  70.681 -67.614  1.00  3.19 C
ATOM    278  O   SER    37A     -52.364  71.585 -68.445  1.00  4.82 O
ATOM    279  N   LEU    38A     -51.322  69.838 -67.499  1.00  2.75 N
ATOM    280  CA  LEU    38A     -50.177  69.927 -68.396  1.00  3.88 C
ATOM    281  CB  LEU    38A     -48.880  69.593 -67.644  1.00  4.62 C
ATOM    282  CG  LEU    38A     -48.507  70.597 -66.544  1.00  5.58 C
ATOM    283  CD1 LEU    38A     -47.239  70.148 -65.826  1.00  6.10 C
ATOM    284  CD2 LEU    38A     -48.292  71.982 -67.165  1.00  5.95 C
ATOM    285  C   LEU    38A     -50.401  68.952 -69.556  1.00  3.84 C
ATOM    286  O   LEU    38A     -50.354  67.739 -69.378  1.00  3.27 O
ATOM    287  N   VAL    39A     -50.650  69.500 -70.741  1.00  4.59 N
ATOM    288  CA  VAL    39A     -50.922  68.705 -71.933  1.00  4.03 C
ATOM    289  CB  VAL    39A     -51.904  69.453 -72.865  1.00  5.84 C
ATOM    290  CG1 VAL    39A     -52.140  68.643 -74.132  1.00  3.93 C
ATOM    291  CG2 VAL    39A     -53.218  69.714 -72.141  1.00  4.78 C
ATOM    292  C   VAL    39A     -49.689  68.319 -72.751  1.00  3.73 C
ATOM    293  O   VAL    39A     -48.865  69.166 -73.099  1.00  4.42 O
ATOM    294  N   ALA    40A     -49.568  67.031 -73.058  1.00  3.86 N
ATOM    295  CA  ALA    40A     -48.451  66.550 -73.859  1.00  2.21 C
ATOM    296  CB  ALA    40A     -48.411  65.017 -73.856  1.00  3.56 C
ATOM    297  C   ALA    40A     -48.657  67.068 -75.272  1.00  4.75 C
ATOM    298  O   ALA    40A     -49.532  66.586 -76.001  1.00  6.49 O
ATOM    299  N   ASP    41A     -47.863  68.063 -75.654  1.00  4.30 N
ATOM    300  CA  ASP    41A     -47.970  68.649 -76.977  1.00  4.66 C
ATOM    301  CB  ASP    41A     -48.323  70.137 -76.872  1.00  2.84 C
ATOM    302  CG  ASP    41A     -48.863  70.693 -78.169  1.00  2.42 C
```

FIGURE 9 - 5

```
ATOM    303  OD1 ASP   41A    -48.415  70.242 -79.237  1.00  4.68 O
ATOM    304  OD2 ASP   41A    -49.732  71.590 -78.126  1.00  3.54 O
ATOM    305  C   ASP   41A    -46.655  68.500 -77.730  1.00  3.54 C
ATOM    306  O   ASP   41A    -45.686  69.200 -77.449  1.00  3.14 O
ATOM    307  N   ARG   42A    -46.634  67.589 -78.693  1.00  4.75 N
ATOM    308  CA  ARG   42A    -45.436  67.345 -79.479  1.00  3.76 C
ATOM    309  CB  ARG   42A    -45.709  66.244 -80.502  1.00  4.61 C
ATOM    310  CG  ARG   42A    -44.490  65.802 -81.283  1.00  6.73 C
ATOM    311  CD  ARG   42A    -44.836  64.660 -82.239  1.00  6.84 C
ATOM    312  NE  ARG   42A    -45.189  63.406 -81.570  1.00  7.48 N
ATOM    313  CZ  ARG   42A    -44.312  62.543 -81.061  1.00  6.12 C
ATOM    314  NH1 ARG   42A    -43.008  62.789 -81.125  1.00  2.39 N
ATOM    315  NH2 ARG   42A    -44.737  61.407 -80.521  1.00  4.79 N
ATOM    316  C   ARG   42A    -44.966  68.613 -80.187  1.00  3.81 C
ATOM    317  O   ARG   42A    -43.766  68.807 -80.393  1.00  3.12 O
ATOM    318  N   LYS   43A    -45.906  69.486 -80.538  1.00  3.21 N
ATOM    319  CA  LYS   43A    -45.549  70.722 -81.227  1.00  2.31 C
ATOM    320  CB  LYS   43A    -46.802  71.518 -81.602  1.00  3.51 C
ATOM    321  CG  LYS   43A    -47.740  70.749 -82.527  1.00  5.83 C
ATOM    322  CD  LYS   43A    -48.846  71.627 -83.102  1.00 10.95 C
ATOM    323  CE  LYS   43A    -49.718  72.252 -82.020  1.00 12.62 C
ATOM    324  NZ  LYS   43A    -49.009  73.316 -81.261  1.00 14.99 N
ATOM    325  C   LYS   43A    -44.592  71.602 -80.438  1.00  2.79 C
ATOM    326  O   LYS   43A    -43.844  72.374 -81.028  1.00  4.45 O
ATOM    327  N   TYR   44A    -44.607  71.497 -79.113  1.00  1.00 N
ATOM    328  CA  TYR   44A    -43.697  72.308 -78.306  1.00  2.04 C
ATOM    329  CB  TYR   44A    -44.488  73.213 -77.367  1.00  1.00 C
ATOM    330  CG  TYR   44A    -45.523  74.024 -78.114  1.00  2.73 C
ATOM    331  CD1 TYR   44A    -46.879  73.736 -77.995  1.00  4.30 C
ATOM    332  CE1 TYR   44A    -47.832  74.440 -78.723  1.00  5.05 C
ATOM    333  CD2 TYR   44A    -45.137  75.043 -78.983  1.00  5.72 C
ATOM    334  CE2 TYR   44A    -46.079  75.756 -79.720  1.00  7.11 C
ATOM    335  CZ  TYR   44A    -47.425  75.450 -79.585  1.00  8.71 C
ATOM    336  OH  TYR   44A    -48.362  76.168 -80.297  1.00 12.07 O
ATOM    337  C   TYR   44A    -42.711  71.450 -77.525  1.00  1.00 C
ATOM    338  O   TYR   44A    -42.105  71.900 -76.557  1.00  1.08 O
ATOM    339  N   PHE   45A    -42.567  70.203 -77.955  1.00  1.59 N
ATOM    340  CA  PHE   45A    -41.618  69.279 -77.342  1.00  1.00 C
ATOM    341  CB  PHE   45A    -40.208  69.675 -77.774  1.00  2.88 C
ATOM    342  CG  PHE   45A    -40.068  69.877 -79.253  1.00  5.61 C
ATOM    343  CD1 PHE   45A    -40.115  68.793 -80.125  1.00  9.29 C
ATOM    344  CD2 PHE   45A    -39.915  71.153 -79.782  1.00  7.48 C
ATOM    345  CE1 PHE   45A    -40.011  68.976 -81.501  1.00  9.29 C
ATOM    346  CE2 PHE   45A    -39.811  71.348 -81.159  1.00  8.29 C
ATOM    347  CZ  PHE   45A    -39.860  70.255 -82.020  1.00  7.53 C
ATOM    348  C   PHE   45A    -41.685  69.214 -75.820  1.00  1.00 C
ATOM    349  O   PHE   45A    -40.650  69.226 -75.149  1.00  1.00 O
ATOM    350  N   GLY   46A    -42.890  69.144 -75.266  1.00  1.00 N
ATOM    351  CA  GLY   46A    -42.997  69.070 -73.821  1.00  1.00 C
ATOM    352  C   GLY   46A    -44.421  69.076 -73.316  1.00  1.01 C
ATOM    353  O   GLY   46A    -45.357  68.859 -74.082  1.00  2.94 O
ATOM    354  N   LEU   47A    -44.576  69.307 -72.016  1.00  2.40 N
ATOM    355  CA  LEU   47A    -45.888  69.365 -71.391  1.00  1.00 C
ATOM    356  CB  LEU   47A    -45.860  68.695 -70.020  1.00  1.16 C
ATOM    357  CG  LEU   47A    -45.365  67.246 -70.021  1.00  2.74 C
ATOM    358  CD1 LEU   47A    -45.407  66.684 -68.612  1.00  4.58 C
ATOM    359  CD2 LEU   47A    -46.236  66.414 -70.952  1.00  1.00 C
ATOM    360  C   LEU   47A    -46.218  70.843 -71.248  1.00  1.26 C
ATOM    361  O   LEU   47A    -45.431  71.604 -70.680  1.00  1.00 O
ATOM    362  N   VAL   48A    -47.379  71.236 -71.762  1.00  1.81 N
ATOM    363  CA  VAL   48A    -47.807  72.635 -71.742  1.00  2.91 C
```

FIGURE 9 - 6

```
ATOM    364  CB  VAL   48A    -47.925  73.187 -73.186  1.00  5.03 C
ATOM    365  CG1 VAL   48A    -46.620  72.969 -73.935  1.00  6.02 C
ATOM    366  CG2 VAL   48A    -49.068  72.501 -73.915  1.00 10.10 C
ATOM    367  C   VAL   48A    -49.140  72.885 -71.041  1.00  2.18 C
ATOM    368  O   VAL   48A    -50.125  72.187 -71.285  1.00  3.39 O
ATOM    369  N   GLY   49A    -49.157  73.894 -70.174  1.00  2.60 N
ATOM    370  CA  GLY   49A    -50.376  74.248 -69.463  1.00  4.21 C
ATOM    371  C   GLY   49A    -50.638  75.740 -69.562  1.00  3.44 C
ATOM    372  O   GLY   49A    -49.707  76.529 -69.735  1.00  2.81 O
ATOM    373  N   LYS   50A    -51.899  76.141 -69.440  1.00  4.77 N
ATOM    374  CA  LYS   50A    -52.245  77.555 -69.538  1.00  6.93 C
ATOM    375  CB  LYS   50A    -52.813  77.847 -70.932  1.00 10.17 C
ATOM    376  CG  LYS   50A    -53.131  79.309 -71.197  1.00 16.75 C
ATOM    377  CD  LYS   50A    -53.654  79.518 -72.617  1.00 20.36 C
ATOM    378  CE  LYS   50A    -55.002  78.842 -72.838  1.00 24.17 C
ATOM    379  NZ  LYS   50A    -56.094  79.494 -72.060  1.00 25.42 N
ATOM    380  C   LYS   50A    -53.252  78.004 -68.490  1.00  5.87 C
ATOM    381  O   LYS   50A    -54.130  77.242 -68.095  1.00  3.41 O
ATOM    382  N   PHE   51A    -53.118  79.250 -68.044  1.00  7.45 N
ATOM    383  CA  PHE   51A    -54.053  79.819 -67.082  1.00  7.81 C
ATOM    384  CB  PHE   51A    -53.714  79.387 -65.647  1.00 10.96 C
ATOM    385  CG  PHE   51A    -52.503  80.052 -65.059  1.00 13.30 C
ATOM    386  CD1 PHE   51A    -52.621  81.249 -64.355  1.00 15.42 C
ATOM    387  CD2 PHE   51A    -51.250  79.456 -65.161  1.00 15.06 C
ATOM    388  CE1 PHE   51A    -51.506  81.837 -63.756  1.00 16.08 C
ATOM    389  CE2 PHE   51A    -50.131  80.034 -64.568  1.00 16.03 C
ATOM    390  CZ  PHE   51A    -50.259  81.228 -63.861  1.00 17.18 C
ATOM    391  C   PHE   51A    -54.031  81.336 -67.234  1.00  7.87 C
ATOM    392  O   PHE   51A    -53.092  81.893 -67.803  1.00  8.80 O
ATOM    393  N   SER   52A    -55.074  81.994 -66.749  1.00  7.15 N
ATOM    394  CA  SER   52A    -55.166  83.441 -66.871  1.00  6.48 C
ATOM    395  CB  SER   52A    -56.603  83.845 -67.218  1.00  9.11 C
ATOM    396  OG  SER   52A    -57.052  83.186 -68.389  1.00 13.76 O
ATOM    397  C   SER   52A    -54.740  84.176 -65.614  1.00  5.40 C
ATOM    398  O   SER   52A    -54.953  83.701 -64.502  1.00  5.52 O
ATOM    399  N   HIS   53A    -54.122  85.337 -65.804  1.00  6.43 N
ATOM    400  CA  HIS   53A    -53.707  86.168 -64.688  1.00  3.93 C
ATOM    401  CB  HIS   53A    -52.314  85.811 -64.180  1.00  4.24 C
ATOM    402  CG  HIS   53A    -51.992  86.456 -62.870  1.00  6.40 C
ATOM    403  CD2 HIS   53A    -52.566  86.324 -61.649  1.00  3.62 C
ATOM    404  ND1 HIS   53A    -51.012  87.414 -62.730  1.00  3.10 N
ATOM    405  CE1 HIS   53A    -50.996  87.845 -61.482  1.00  3.97 C
ATOM    406  NE2 HIS   53A    -51.929  87.200 -60.805  1.00  4.87 N
ATOM    407  C   HIS   53A    -53.720  87.620 -65.110  1.00  5.30 C
ATOM    408  O   HIS   53A    -52.971  88.032 -66.001  1.00  2.27 O
ATOM    409  N   GLN   54A    -54.586  88.386 -64.457  1.00  4.15 N
ATOM    410  CA  GLN   54A    -54.741  89.799 -64.740  1.00  6.05 C
ATOM    411  CB  GLN   54A    -53.488  90.570 -64.276  1.00  5.48 C
ATOM    412  CG  GLN   54A    -53.287  90.479 -62.746  1.00  2.51 C
ATOM    413  CD  GLN   54A    -52.196  91.385 -62.184  1.00  6.58 C
ATOM    414  OE1 GLN   54A    -52.023  91.473 -60.965  1.00  5.86 O
ATOM    415  NE2 GLN   54A    -51.458  92.059 -63.060  1.00  5.07 N
ATOM    416  C   GLN   54A    -55.044  90.028 -66.224  1.00  7.00 C
ATOM    417  O   GLN   54A    -54.496  90.932 -66.860  1.00  7.88 O
ATOM    418  N   GLY   55A    -55.921  89.181 -66.760  1.00  6.86 N
ATOM    419  CA  GLY   55A    -56.349  89.293 -68.145  1.00  8.85 C
ATOM    420  C   GLY   55A    -55.434  88.773 -69.235  1.00  9.69 C
ATOM    421  O   GLY   55A    -55.723  88.955 -70.414  1.00  9.96 O
ATOM    422  N   ALA   56A    -54.335  88.129 -68.860  1.00  9.11 N
ATOM    423  CA  ALA   56A    -53.404  87.612 -69.854  1.00  8.72 C
ATOM    424  CB  ALA   56A    -52.103  88.402 -69.820  1.00  8.99 C
```

FIGURE 9 - 7

```
ATOM    425  C   ALA   56A    -53.119  86.135 -69.639  1.00  9.60 C
ATOM    426  O   ALA   56A    -53.235  85.620 -68.524  1.00  9.28 O
ATOM    427  N   ASP   57A    -52.754  85.461 -70.725  1.00  8.89 N
ATOM    428  CA  ASP   57A    -52.435  84.042 -70.688  1.00  8.00 C
ATOM    429  CB  ASP   57A    -52.532  83.438 -72.092  1.00 10.97 C
ATOM    430  CG  ASP   57A    -53.954  83.331 -72.589  1.00 14.23 C
ATOM    431  OD1 ASP   57A    -54.137  83.103 -73.803  1.00 17.50 O
ATOM    432  OD2 ASP   57A    -54.885  83.457 -71.768  1.00 15.79 O
ATOM    433  C   ASP   57A    -51.023  83.830 -70.170  1.00  5.94 C
ATOM    434  O   ASP   57A    -50.082  84.480 -70.631  1.00  7.57 O
ATOM    435  N   VAL   58A    -50.892  82.937 -69.196  1.00  3.48 N
ATOM    436  CA  VAL   58A    -49.598  82.575 -68.631  1.00  3.89 C
ATOM    437  CB  VAL   58A    -49.552  82.833 -67.110  1.00  4.00 C
ATOM    438  CG1 VAL   58A    -48.184  82.454 -66.550  1.00  4.48 C
ATOM    439  CG2 VAL   58A    -49.831  84.313 -66.834  1.00  3.15 C
ATOM    440  C   VAL   58A    -49.455  81.079 -68.907  1.00  4.13 C
ATOM    441  O   VAL   58A    -50.336  80.289 -68.559  1.00  5.53 O
ATOM    442  N   ARG   59A    -48.359  80.686 -69.540  1.00  2.93 N
ATOM    443  CA  ARG   59A    -48.151  79.277 -69.863  1.00  3.40 C
ATOM    444  CB  ARG   59A    -47.866  79.124 -71.361  1.00  3.15 C
ATOM    445  CG  ARG   59A    -49.086  79.427 -72.239  1.00  8.12 C
ATOM    446  CD  ARG   59A    -48.850  79.097 -73.707  1.00 10.41 C
ATOM    447  NE  ARG   59A    -50.061  79.248 -74.515  1.00 14.03 N
ATOM    448  CZ  ARG   59A    -50.678  80.405 -74.750  1.00 16.18 C
ATOM    449  NH1 ARG   59A    -51.773  80.434 -75.497  1.00 17.25 N
ATOM    450  NH2 ARG   59A    -50.204  81.536 -74.246  1.00 16.61 N
ATOM    451  C   ARG   59A    -47.056  78.596 -69.050  1.00  2.36 C
ATOM    452  O   ARG   59A    -46.051  79.210 -68.696  1.00  2.33 O
ATOM    453  N   LEU   60A    -47.288  77.326 -68.741  1.00  2.20 N
ATOM    454  CA  LEU   60A    -46.345  76.508 -67.985  1.00  2.36 C
ATOM    455  CB  LEU   60A    -47.054  75.796 -66.833  1.00  1.24 C
ATOM    456  CG  LEU   60A    -47.845  76.682 -65.869  1.00  1.00 C
ATOM    457  CD1 LEU   60A    -48.381  75.832 -64.741  1.00  3.70 C
ATOM    458  CD2 LEU   60A    -46.944  77.782 -65.309  1.00  1.97 C
ATOM    459  C   LEU   60A    -45.780  75.476 -68.950  1.00  1.70 C
ATOM    460  O   LEU   60A    -46.518  74.881 -69.734  1.00  2.01 O
ATOM    461  N   LEU   61A    -44.476  75.256 -68.886  1.00  1.00 N
ATOM    462  CA  LEU   61A    -43.842  74.315 -69.787  1.00  1.70 C
ATOM    463  CB  LEU   61A    -43.191  75.082 -70.940  1.00  2.53 C
ATOM    464  CG  LEU   61A    -42.246  74.326 -71.874  1.00  1.00 C
ATOM    465  CD1 LEU   61A    -43.024  73.279 -72.660  1.00  2.50 C
ATOM    466  CD2 LEU   61A    -41.562  75.316 -72.827  1.00  4.69 C
ATOM    467  C   LEU   61A    -42.801  73.422 -69.135  1.00  1.00 C
ATOM    468  O   LEU   61A    -41.926  73.884 -68.408  1.00  1.00 O
ATOM    469  N   ILE   62A    -42.922  72.125 -69.381  1.00  1.63 N
ATOM    470  CA  ILE   62A    -41.940  71.178 -68.881  1.00  2.33 C
ATOM    471  CB  ILE   62A    -42.536  70.133 -67.912  1.00  2.79 C
ATOM    472  CG2 ILE   62A    -41.476  69.067 -67.589  1.00  1.44 C
ATOM    473  CG1 ILE   62A    -43.007  70.821 -66.630  1.00  4.70 C
ATOM    474  CD1 ILE   62A    -43.537  69.867 -65.572  1.00  3.15 C
ATOM    475  C   ILE   62A    -41.481  70.479 -70.140  1.00  3.04 C
ATOM    476  O   ILE   62A    -42.198  69.643 -70.686  1.00  3.83 O
ATOM    477  N   PRO   63A    -40.303  70.858 -70.656  1.00  3.51 N
ATOM    478  CD  PRO   63A    -39.399  71.924 -70.201  1.00  2.92 C
ATOM    479  CA  PRO   63A    -39.787  70.226 -71.871  1.00  2.01 C
ATOM    480  CB  PRO   63A    -38.440  70.913 -72.073  1.00  2.96 C
ATOM    481  CG  PRO   63A    -38.642  72.250 -71.465  1.00  3.76 C
ATOM    482  C   PRO   63A    -39.616  68.734 -71.636  1.00  2.67 C
ATOM    483  O   PRO   63A    -39.275  68.314 -70.528  1.00  2.30 O
ATOM    484  N   THR   64A    -39.835  67.945 -72.682  1.00  1.29 N
ATOM    485  CA  THR   64A    -39.692  66.500 -72.589  1.00  2.46 C
```

FIGURE 9 - 8

```
ATOM    486  CB  THR   64A     -40.997  65.779 -72.991  1.00  3.31 C
ATOM    487  OG1 THR   64A     -41.372  66.160 -74.320  1.00  5.54 O
ATOM    488  CG2 THR   64A     -42.117  66.156 -72.022  1.00  3.62 C
ATOM    489  C   THR   64A     -38.543  66.038 -73.474  1.00  2.95 C
ATOM    490  O   THR   64A     -38.408  64.855 -73.785  1.00  1.35 O
ATOM    491  N   THR   65A     -37.732  67.001 -73.895  1.00  2.35 N
ATOM    492  CA  THR   65A     -36.552  66.720 -74.688  1.00  3.31 C
ATOM    493  CB  THR   65A     -36.014  67.981 -75.387  1.00  1.69 C
ATOM    494  OG1 THR   65A     -35.718  68.977 -74.395  1.00  3.42 O
ATOM    495  CG2 THR   65A     -37.029  68.542 -76.371  1.00  1.00 C
ATOM    496  C   THR   65A     -35.576  66.404 -73.576  1.00  4.91 C
ATOM    497  O   THR   65A     -35.908  66.541 -72.395  1.00  5.83 O
ATOM    498  N   TYR   66A     -34.377  65.972 -73.922  1.00  8.39 N
ATOM    499  CA  TYR   66A     -33.420  65.734 -72.868  1.00  8.97 C
ATOM    500  CB  TYR   66A     -32.265  64.884 -73.379  1.00 10.94 C
ATOM    501  CG  TYR   66A     -32.547  63.428 -73.105  1.00 13.20 C
ATOM    502  CD1 TYR   66A     -32.564  62.485 -74.129  1.00 12.97 C
ATOM    503  CE1 TYR   66A     -32.875  61.153 -73.868  1.00 11.59 C
ATOM    504  CD2 TYR   66A     -32.842  63.006 -71.809  1.00 15.12 C
ATOM    505  CE2 TYR   66A     -33.148  61.684 -71.534  1.00 16.61 C
ATOM    506  CZ  TYR   66A     -33.165  60.763 -72.564  1.00 15.75 C
ATOM    507  OH  TYR   66A     -33.471  59.461 -72.272  1.00 14.48 O
ATOM    508  C   TYR   66A     -32.983  67.115 -72.419  1.00  6.85 C
ATOM    509  O   TYR   66A     -33.195  68.084 -73.140  1.00  7.38 O
ATOM    510  N   MET   67A     -32.411  67.217 -71.225  1.00  7.40 N
ATOM    511  CA  MET   67A     -31.991  68.515 -70.706  1.00  7.99 C
ATOM    512  CB  MET   67A     -31.209  68.340 -69.401  1.00  9.03 C
ATOM    513  CG  MET   67A     -30.887  69.650 -68.687  1.00  5.11 C
ATOM    514  SD  MET   67A     -32.369  70.459 -68.017  1.00 11.05 S
ATOM    515  CE  MET   67A     -32.474  69.674 -66.396  1.00  6.71 C
ATOM    516  C   MET   67A     -31.151  69.332 -71.687  1.00  8.16 C
ATOM    517  O   MET   67A     -31.331  70.547 -71.802  1.00  4.96 O
ATOM    518  N   ASN   68A     -30.235  68.671 -72.392  1.00  9.23 N
ATOM    519  CA  ASN   68A     -29.358  69.364 -73.338  1.00 11.14 C
ATOM    520  CB  ASN   68A     -28.192  68.446 -73.739  1.00 16.37 C
ATOM    521  CG  ASN   68A     -28.638  67.039 -74.078  1.00 19.09 C
ATOM    522  OD1 ASN   68A     -27.837  66.101 -74.054  1.00 22.31 O
ATOM    523  ND2 ASN   68A     -29.912  66.881 -74.404  1.00 22.11 N
ATOM    524  C   ASN   68A     -30.030  69.946 -74.589  1.00  9.74 C
ATOM    525  O   ASN   68A     -29.389  70.662 -75.366  1.00  9.29 O
ATOM    526  N   ARG   69A     -31.319  69.663 -74.770  1.00  8.50 N
ATOM    527  CA  ARG   69A     -32.061  70.176 -75.920  1.00  7.23 C
ATOM    528  CB  ARG   69A     -32.507  69.018 -76.812  1.00 10.71 C
ATOM    529  CG  ARG   69A     -31.351  68.250 -77.435  1.00 13.81 C
ATOM    530  CD  ARG   69A     -30.673  69.057 -78.529  1.00 15.95 C
ATOM    531  NE  ARG   69A     -29.421  68.432 -78.947  1.00 19.05 N
ATOM    532  CZ  ARG   69A     -28.663  68.862 -79.949  1.00 20.33 C
ATOM    533  NH1 ARG   69A     -27.536  68.228 -80.247  1.00 19.01 N
ATOM    534  NH2 ARG   69A     -29.037  69.918 -80.659  1.00 23.40 N
ATOM    535  C   ARG   69A     -33.281  70.995 -75.490  1.00  6.80 C
ATOM    536  O   ARG   69A     -34.140  71.326 -76.307  1.00  1.63 O
ATOM    537  N   SER   70A     -33.332  71.322 -74.201  1.00  4.11 N
ATOM    538  CA  SER   70A     -34.431  72.093 -73.608  1.00  4.52 C
ATOM    539  CB  SER   70A     -34.078  72.462 -72.162  1.00  5.63 C
ATOM    540  OG  SER   70A     -33.981  71.315 -71.349  1.00 11.27 O
ATOM    541  C   SER   70A     -34.792  73.374 -74.368  1.00  4.30 C
ATOM    542  O   SER   70A     -35.961  73.777 -74.399  1.00  5.52 O
ATOM    543  N   GLY   71A     -33.786  74.012 -74.963  1.00  3.88 N
ATOM    544  CA  GLY   71A     -34.005  75.253 -75.691  1.00  2.13 C
ATOM    545  C   GLY   71A     -34.912  75.099 -76.897  1.00  2.45 C
ATOM    546  O   GLY   71A     -35.530  76.055 -77.360  1.00  1.26 O
```

FIGURE 9 -9

```
ATOM    547  N   GLN   72A    -34.971  73.880 -77.414  1.00  2.29 N
ATOM    548  CA  GLN   72A    -35.802  73.556 -78.557  1.00  5.21 C
ATOM    549  CB  GLN   72A    -35.507  72.110 -78.958  1.00 10.57 C
ATOM    550  CG  GLN   72A    -36.235  71.574 -80.159  1.00 15.03 C
ATOM    551  CD  GLN   72A    -35.755  70.177 -80.499  1.00 18.58 C
ATOM    552  OE1 GLN   72A    -34.558  69.958 -80.694  1.00 20.08 O
ATOM    553  NE2 GLN   72A    -36.678  69.225 -80.561  1.00 18.74 N
ATOM    554  C   GLN   72A    -37.268  73.716 -78.143  1.00  4.30 C
ATOM    555  O   GLN   72A    -38.101  74.244 -78.893  1.00  1.00 O
ATOM    556  N   SER   73A    -37.576  73.257 -76.934  1.00  2.01 N
ATOM    557  CA  SER   73A    -38.937  73.339 -76.414  1.00  2.93 C
ATOM    558  CB  SER   73A    -39.057  72.492 -75.139  1.00  4.70 C
ATOM    559  OG  SER   73A    -40.375  72.520 -74.623  1.00  2.80 O
ATOM    560  C   SER   73A    -39.327  74.786 -76.120  1.00  2.58 C
ATOM    561  O   SER   73A    -40.367  75.270 -76.575  1.00  1.00 O
ATOM    562  N   VAL   74A    -38.478  75.478 -75.370  1.00  1.77 N
ATOM    563  CA  VAL   74A    -38.732  76.867 -75.009  1.00  2.80 C
ATOM    564  CB  VAL   74A    -37.571  77.430 -74.160  1.00  2.28 C
ATOM    565  CG1 VAL   74A    -37.823  78.886 -73.825  1.00  1.00 C
ATOM    566  CG2 VAL   74A    -37.431  76.613 -72.870  1.00  1.62 C
ATOM    567  C   VAL   74A    -38.953  77.767 -76.225  1.00  2.25 C
ATOM    568  O   VAL   74A    -39.902  78.554 -76.258  1.00  1.00 O
ATOM    569  N   ALA   75A    -38.093  77.640 -77.230  1.00  2.77 N
ATOM    570  CA  ALA   75A    -38.215  78.469 -78.425  1.00  3.35 C
ATOM    571  CB  ALA   75A    -36.984  78.309 -79.299  1.00  1.00 C
ATOM    572  C   ALA   75A    -39.479  78.162 -79.232  1.00  1.73 C
ATOM    573  O   ALA   75A    -40.117  79.070 -79.759  1.00  4.44 O
ATOM    574  N   ALA   76A    -39.840  76.890 -79.332  1.00  3.55 N
ATOM    575  CA  ALA   76A    -41.034  76.501 -80.077  1.00  2.93 C
ATOM    576  CB  ALA   76A    -41.199  74.971 -80.052  1.00  1.00 C
ATOM    577  C   ALA   76A    -42.280  77.156 -79.491  1.00  3.63 C
ATOM    578  O   ALA   76A    -43.140  77.650 -80.219  1.00  3.65 O
ATOM    579  N   LEU   77A    -42.385  77.142 -78.168  1.00  4.48 N
ATOM    580  CA  LEU   77A    -43.551  77.719 -77.517  1.00  5.34 C
ATOM    581  CB  LEU   77A    -43.714  77.150 -76.103  1.00  6.17 C
ATOM    582  CG  LEU   77A    -44.995  77.554 -75.359  1.00  5.07 C
ATOM    583  CD1 LEU   77A    -46.220  77.059 -76.122  1.00  6.80 C
ATOM    584  CD2 LEU   77A    -44.982  76.956 -73.962  1.00  7.62 C
ATOM    585  C   LEU   77A    -43.448  79.234 -77.460  1.00  5.10 C
ATOM    586  O   LEU   77A    -44.408  79.942 -77.759  1.00  5.93 O
ATOM    587  N   ALA   78A    -42.275  79.728 -77.078  1.00  5.02 N
ATOM    588  CA  ALA   78A    -42.058  81.164 -76.980  1.00  5.59 C
ATOM    589  CB  ALA   78A    -40.630  81.438 -76.531  1.00  5.72 C
ATOM    590  C   ALA   78A    -42.330  81.851 -78.316  1.00  6.96 C
ATOM    591  O   ALA   78A    -43.078  82.834 -78.386  1.00  5.88 O
ATOM    592  N   GLY   79A    -41.726  81.327 -79.376  1.00  4.76 N
ATOM    593  CA  GLY   79A    -41.906  81.921 -80.688  1.00  3.15 C
ATOM    594  C   GLY   79A    -43.329  81.899 -81.201  1.00  2.65 C
ATOM    595  O   GLY   79A    -43.815  82.887 -81.766  1.00  2.74 O
ATOM    596  N   PHE   80A    -44.001  80.774 -80.999  1.00  2.79 N
ATOM    597  CA  PHE   80A    -45.376  80.593 -81.454  1.00  2.15 C
ATOM    598  CB  PHE   80A    -45.870  79.196 -81.073  1.00  3.34 C
ATOM    599  CG  PHE   80A    -47.173  78.817 -81.718  1.00  3.75 C
ATOM    600  CD1 PHE   80A    -47.188  78.149 -82.940  1.00  4.24 C
ATOM    601  CD2 PHE   80A    -48.385  79.140 -81.111  1.00  5.10 C
ATOM    602  CE1 PHE   80A    -48.393  77.806 -83.552  1.00  5.43 C
ATOM    603  CE2 PHE   80A    -49.599  78.802 -81.713  1.00  5.78 C
ATOM    604  CZ  PHE   80A    -49.603  78.133 -82.935  1.00  6.24 C
ATOM    605  C   PHE   80A    -46.328  81.641 -80.881  1.00  2.81 C
ATOM    606  O   PHE   80A    -47.224  82.118 -81.577  1.00  2.29 O
ATOM    607  N   PHE   81A    -46.134  81.993 -79.613  1.00  3.22 N
```

FIGURE 9 - 10

```
ATOM    608  CA  PHE   81A    -46.986  82.985 -78.959  1.00  5.90 C
ATOM    609  CB  PHE   81A    -47.438  82.458 -77.591  1.00  4.24 C
ATOM    610  CG  PHE   81A    -48.372  81.285 -77.686  1.00  6.23 C
ATOM    611  CD1 PHE   81A    -49.632  81.433 -78.256  1.00  7.74 C
ATOM    612  CD2 PHE   81A    -47.977  80.019 -77.260  1.00  6.43 C
ATOM    613  CE1 PHE   81A    -50.484  80.343 -78.406  1.00  8.95 C
ATOM    614  CE2 PHE   81A    -48.821  78.920 -77.407  1.00  7.65 C
ATOM    615  CZ  PHE   81A    -50.079  79.082 -77.982  1.00  6.76 C
ATOM    616  C   PHE   81A    -46.321  84.353 -78.815  1.00  7.85 C
ATOM    617  O   PHE   81A    -46.822  85.221 -78.104  1.00  8.35 O
ATOM    618  N   ARG   82A    -45.196  84.532 -79.502  1.00  9.48 N
ATOM    619  CA  ARG   82A    -44.444  85.784 -79.485  1.00 13.08 C
ATOM    620  CB  ARG   82A    -45.196  86.862 -80.282  1.00 17.88 C
ATOM    621  CG  ARG   82A    -45.189  86.625 -81.794  1.00 24.42 C
ATOM    622  CD  ARG   82A    -45.897  85.328 -82.157  1.00 31.34 C
ATOM    623  NE  ARG   82A    -45.614  84.884 -83.521  1.00 37.09 N
ATOM    624  CZ  ARG   82A    -46.002  85.522 -84.619  1.00 39.80 C
ATOM    625  NH1 ARG   82A    -45.694  85.032 -85.812  1.00 41.45 N
ATOM    626  NH2 ARG   82A    -46.699  86.647 -84.529  1.00 41.70 N
ATOM    627  C   ARG   82A    -44.106  86.302 -78.086  1.00 12.16 C
ATOM    628  O   ARG   82A    -44.303  87.480 -77.773  1.00 11.28 O
ATOM    629  N   ILE   83A    -43.582  85.412 -77.251  1.00  9.21 N
ATOM    630  CA  ILE   83A    -43.197  85.770 -75.893  1.00  6.42 C
ATOM    631  CB  ILE   83A    -43.472  84.593 -74.930  1.00  4.96 C
ATOM    632  CG2 ILE   83A    -43.021  84.945 -73.512  1.00  3.65 C
ATOM    633  CG1 ILE   83A    -44.964  84.244 -74.967  1.00  5.53 C
ATOM    634  CD1 ILE   83A    -45.319  82.973 -74.218  1.00  4.43 C
ATOM    635  C   ILE   83A    -41.707  86.131 -75.873  1.00  6.52 C
ATOM    636  O   ILE   83A    -40.860  85.332 -76.274  1.00  4.43 O
ATOM    637  N   ALA   84A    -41.396  87.345 -75.423  1.00  5.32 N
ATOM    638  CA  ALA   84A    -40.011  87.802 -75.361  1.00  4.65 C
ATOM    639  CB  ALA   84A    -39.969  89.326 -75.336  1.00  6.43 C
ATOM    640  C   ALA   84A    -39.297  87.241 -74.133  1.00  4.51 C
ATOM    641  O   ALA   84A    -39.928  86.930 -73.131  1.00  4.29 O
ATOM    642  N   PRO   85A    -37.961  87.113 -74.197  1.00  4.93 N
ATOM    643  CD  PRO   85A    -37.083  87.425 -75.340  1.00  5.99 C
ATOM    644  CA  PRO   85A    -37.178  86.589 -73.075  1.00  5.44 C
ATOM    645  CB  PRO   85A    -35.739  86.829 -73.525  1.00  7.50 C
ATOM    646  CG  PRO   85A    -35.831  86.633 -75.008  1.00  6.24 C
ATOM    647  C   PRO   85A    -37.486  87.259 -71.733  1.00  6.13 C
ATOM    648  O   PRO   85A    -37.449  86.606 -70.689  1.00  4.55 O
ATOM    649  N   ASP   86A    -37.793  88.556 -71.758  1.00  3.98 N
ATOM    650  CA  ASP   86A    -38.093  89.277 -70.523  1.00  5.85 C
ATOM    651  CB  ASP   86A    -38.263  90.773 -70.794  1.00  8.63 C
ATOM    652  CG  ASP   86A    -38.162  91.611 -69.531  1.00 14.35 C
ATOM    653  OD1 ASP   86A    -38.734  92.722 -69.503  1.00 16.61 O
ATOM    654  OD2 ASP   86A    -37.495  91.175 -68.569  1.00 16.31 O
ATOM    655  C   ASP   86A    -39.358  88.757 -69.847  1.00  3.15 C
ATOM    656  O   ASP   86A    -39.566  88.985 -68.659  1.00  2.69 O
ATOM    657  N   ALA   87A    -40.206  88.072 -70.605  1.00  3.41 N
ATOM    658  CA  ALA   87A    -41.445  87.541 -70.047  1.00  3.21 C
ATOM    659  CB  ALA   87A    -42.605  87.779 -71.030  1.00  3.94 C
ATOM    660  C   ALA   87A    -41.340  86.056 -69.701  1.00  2.99 C
ATOM    661  O   ALA   87A    -42.348  85.405 -69.444  1.00  3.09 O
ATOM    662  N   ILE   88A    -40.120  85.527 -69.689  1.00  2.69 N
ATOM    663  CA  ILE   88A    -39.908  84.118 -69.372  1.00  2.29 C
ATOM    664  CB  ILE   88A    -39.072  83.405 -70.460  1.00  1.74 C
ATOM    665  CG2 ILE   88A    -38.926  81.926 -70.095  1.00  1.00 C
ATOM    666  CG1 ILE   88A    -39.727  83.579 -71.833  1.00  1.88 C
ATOM    667  CD1 ILE   88A    -38.896  83.027 -72.989  1.00  2.05 C
ATOM    668  C   ILE   88A    -39.172  83.903 -68.051  1.00  1.00 C
```

FIGURE 9 - 11

```
ATOM    669  O   ILE   88A    -38.218  84.608 -67.737  1.00  4.48 O
ATOM    670  N   LEU   89A    -39.630  82.924 -67.282  1.00  2.39 N
ATOM    671  CA  LEU   89A    -39.002  82.554 -66.014  1.00  1.56 C
ATOM    672  CB  LEU   89A    -39.986  82.725 -64.848  1.00  3.17 C
ATOM    673  CG  LEU   89A    -39.570  82.166 -63.483  1.00  2.35 C
ATOM    674  CD1 LEU   89A    -38.304  82.856 -62.990  1.00  2.47 C
ATOM    675  CD2 LEU   89A    -40.709  82.365 -62.483  1.00  3.18 C
ATOM    676  C   LEU   89A    -38.638  81.084 -66.162  1.00  2.40 C
ATOM    677  O   LEU   89A    -39.467  80.291 -66.597  1.00  1.00 O
ATOM    678  N   VAL   90A    -37.397  80.727 -65.836  1.00  1.68 N
ATOM    679  CA  VAL   90A    -36.966  79.339 -65.936  1.00  1.00 C
ATOM    680  CB  VAL   90A    -35.786  79.176 -66.925  1.00  2.68 C
ATOM    681  CG1 VAL   90A    -35.411  77.706 -67.039  1.00  1.00 C
ATOM    682  CG2 VAL   90A    -36.152  79.738 -68.290  1.00  4.02 C
ATOM    683  C   VAL   90A    -36.523  78.814 -64.575  1.00  1.00 C
ATOM    684  O   VAL   90A    -35.617  79.363 -63.959  1.00  1.04 O
ATOM    685  N   ALA   91A    -37.170  77.753 -64.108  1.00  1.18 N
ATOM    686  CA  ALA   91A    -36.819  77.157 -62.829  1.00  2.29 C
ATOM    687  CB  ALA   91A    -38.073  76.610 -62.123  1.00  1.00 C
ATOM    688  C   ALA   91A    -35.840  76.033 -63.101  1.00  1.00 C
ATOM    689  O   ALA   91A    -36.057  75.198 -63.978  1.00  1.04 O
ATOM    690  N   HIS   92A    -34.755  76.006 -62.344  1.00  1.00 N
ATOM    691  CA  HIS   92A    -33.757  74.976 -62.542  1.00  2.87 C
ATOM    692  CB  HIS   92A    -32.738  75.429 -63.589  1.00  1.00 C
ATOM    693  CG  HIS   92A    -31.843  76.535 -63.118  1.00  1.50 C
ATOM    694  CD2 HIS   92A    -30.662  76.499 -62.457  1.00  1.00 C
ATOM    695  ND1 HIS   92A    -32.154  77.869 -63.270  1.00  3.55 N
ATOM    696  CE1 HIS   92A    -31.203  78.607 -62.724  1.00  1.00 C
ATOM    697  NE2 HIS   92A    -30.287  77.799 -62.223  1.00  3.77 N
ATOM    698  C   HIS   92A    -33.036  74.677 -61.243  1.00  2.36 C
ATOM    699  O   HIS   92A    -33.082  75.478 -60.301  1.00  1.78 O
ATOM    700  N   ASP   93A    -32.373  73.522 -61.207  1.00  1.80 N
ATOM    701  CA  ASP   93A    -31.610  73.102 -60.036  1.00  1.22 C
ATOM    702  CB  ASP   93A    -31.261  71.604 -60.113  1.00  1.00 C
ATOM    703  CG  ASP   93A    -30.614  71.204 -61.438  1.00  2.03 C
ATOM    704  OD1 ASP   93A    -31.238  71.393 -62.497  1.00  4.82 O
ATOM    705  OD2 ASP   93A    -29.478  70.684 -61.423  1.00  2.22 O
ATOM    706  C   ASP   93A    -30.340  73.939 -59.946  1.00  1.11 C
ATOM    707  O   ASP   93A    -29.736  74.286 -60.966  1.00  1.00 O
ATOM    708  N   GLU   94A    -29.933  74.267 -58.725  1.00  1.89 N
ATOM    709  CA  GLU   94A    -28.747  75.095 -58.540  1.00  1.00 C
ATOM    710  CB  GLU   94A    -29.150  76.478 -58.006  1.00  1.00 C
ATOM    711  CG  GLU   94A    -27.963  77.399 -57.699  1.00  1.00 C
ATOM    712  CD  GLU   94A    -27.097  77.656 -58.912  1.00  4.34 C
ATOM    713  OE1 GLU   94A    -27.569  78.321 -59.859  1.00  2.87 O
ATOM    714  OE2 GLU   94A    -25.937  77.185 -58.923  1.00  6.65 O
ATOM    715  C   GLU   94A    -27.730  74.454 -57.614  1.00  1.02 C
ATOM    716  O   GLU   94A    -27.988  74.263 -56.425  1.00  1.00 O
ATOM    717  N   LEU   95A    -26.570  74.136 -58.175  1.00  1.00 N
ATOM    718  CA  LEU   95A    -25.494  73.509 -57.423  1.00  2.70 C
ATOM    719  CB  LEU   95A    -24.397  73.028 -58.387  1.00  2.52 C
ATOM    720  CG  LEU   95A    -24.745  71.889 -59.360  1.00  1.00 C
ATOM    721  CD1 LEU   95A    -23.758  71.878 -60.527  1.00  1.99 C
ATOM    722  CD2 LEU   95A    -24.729  70.555 -58.630  1.00  1.00 C
ATOM    723  C   LEU   95A    -24.878  74.432 -56.369  1.00  2.71 C
ATOM    724  O   LEU   95A    -24.463  73.962 -55.306  1.00  2.36 O
ATOM    725  N   ASP   96A    -24.838  75.735 -56.648  1.00  2.51 N
ATOM    726  CA  ASP   96A    -24.220  76.692 -55.723  1.00  2.80 C
ATOM    727  CB  ASP   96A    -23.729  77.921 -56.491  1.00  3.18 C
ATOM    728  CG  ASP   96A    -22.528  77.616 -57.364  1.00  5.63 C
ATOM    729  OD1 ASP   96A    -21.889  76.560 -57.156  1.00  7.10 O
```

FIGURE 9 - 12

```
ATOM    730  OD2 ASP    96A    -22.213  78.433 -58.251  1.00  6.78 O
ATOM    731  C   ASP    96A    -25.014  77.133 -54.498  1.00  3.41 C
ATOM    732  O   ASP    96A    -24.562  77.994 -53.736  1.00  1.00 O
ATOM    733  N   MET    97A    -26.189  76.543 -54.311  1.00  1.00 N
ATOM    734  CA  MET    97A    -27.033  76.829 -53.160  1.00  2.15 C
ATOM    735  CB  MET    97A    -28.277  77.608 -53.586  1.00  2.06 C
ATOM    736  CG  MET    97A    -27.976  78.995 -54.111  1.00  5.02 C
ATOM    737  SD  MET    97A    -29.393  79.771 -54.919  1.00 11.41 S
ATOM    738  CE  MET    97A    -28.534  80.768 -56.133  1.00  7.74 C
ATOM    739  C   MET    97A    -27.422  75.467 -52.592  1.00  3.33 C
ATOM    740  O   MET    97A    -27.658  74.524 -53.344  1.00  3.24 O
ATOM    741  N   PRO    98A    -27.479  75.338 -51.260  1.00  2.62 N
ATOM    742  CD  PRO    98A    -27.055  76.285 -50.208  1.00  4.54 C
ATOM    743  CA  PRO    98A    -27.844  74.041 -50.686  1.00  3.39 C
ATOM    744  CB  PRO    98A    -27.249  74.113 -49.285  1.00  1.00 C
ATOM    745  CG  PRO    98A    -27.470  75.563 -48.927  1.00  1.91 C
ATOM    746  C   PRO    98A    -29.348  73.792 -50.645  1.00  1.62 C
ATOM    747  O   PRO    98A    -30.154  74.704 -50.864  1.00  3.59 O
ATOM    748  N   PRO    99A    -29.746  72.544 -50.368  1.00  5.30 N
ATOM    749  CD  PRO    99A    -28.937  71.320 -50.230  1.00  2.20 C
ATOM    750  CA  PRO    99A    -31.171  72.231 -50.299  1.00  4.15 C
ATOM    751  CB  PRO    99A    -31.175  70.773 -49.852  1.00  4.44 C
ATOM    752  CG  PRO    99A    -29.957  70.232 -50.508  1.00  5.09 C
ATOM    753  C   PRO    99A    -31.826  73.149 -49.267  1.00  6.14 C
ATOM    754  O   PRO    99A    -31.299  73.323 -48.167  1.00  8.10 O
ATOM    755  N   GLY   100A    -32.965  73.733 -49.624  1.00  5.74 N
ATOM    756  CA  GLY   100A    -33.659  74.621 -48.709  1.00  5.37 C
ATOM    757  C   GLY   100A    -33.565  76.076 -49.118  1.00  7.00 C
ATOM    758  O   GLY   100A    -34.282  76.928 -48.593  1.00  6.78 O
ATOM    759  N   VAL   101A    -32.688  76.361 -50.073  1.00  6.83 N
ATOM    760  CA  VAL   101A    -32.494  77.723 -50.549  1.00  8.14 C
ATOM    761  CB  VAL   101A    -31.007  78.121 -50.455  1.00  7.01 C
ATOM    762  CG1 VAL   101A    -30.789  79.503 -51.071  1.00  6.16 C
ATOM    763  CG2 VAL   101A    -30.571  78.115 -48.995  1.00  6.74 C
ATOM    764  C   VAL   101A    -32.977  77.944 -51.982  1.00  8.68 C
ATOM    765  O   VAL   101A    -32.639  77.183 -52.891  1.00  9.42 O
ATOM    766  N   ALA   102A    -33.769  78.995 -52.170  1.00  6.76 N
ATOM    767  CA  ALA   102A    -34.293  79.343 -53.483  1.00  6.36 C
ATOM    768  CB  ALA   102A    -35.790  79.058 -53.549  1.00  7.15 C
ATOM    769  C   ALA   102A    -34.030  80.821 -53.750  1.00  6.69 C
ATOM    770  O   ALA   102A    -34.325  81.671 -52.914  1.00  5.32 O
ATOM    771  N   LYS   103A    -33.452  81.123 -54.906  1.00  5.55 N
ATOM    772  CA  LYS   103A    -33.182  82.507 -55.252  1.00  5.58 C
ATOM    773  CB  LYS   103A    -31.718  82.854 -54.978  1.00  9.03 C
ATOM    774  CG  LYS   103A    -31.359  82.756 -53.508  1.00 11.36 C
ATOM    775  CD  LYS   103A    -30.098  83.539 -53.181  1.00 13.50 C
ATOM    776  CE  LYS   103A    -29.613  83.226 -51.771  1.00 16.98 C
ATOM    777  NZ  LYS   103A    -30.674  83.450 -50.745  1.00 17.15 N
ATOM    778  C   LYS   103A    -33.548  82.825 -56.694  1.00  5.53 C
ATOM    779  O   LYS   103A    -33.606  81.937 -57.548  1.00  3.59 O
ATOM    780  N   LEU   104A    -33.823  84.101 -56.947  1.00  3.31 N
ATOM    781  CA  LEU   104A    -34.197  84.565 -58.273  1.00  3.87 C
ATOM    782  CB  LEU   104A    -35.455  85.442 -58.190  1.00  3.81 C
ATOM    783  CG  LEU   104A    -36.766  84.741 -57.822  1.00  1.92 C
ATOM    784  CD1 LEU   104A    -37.811  85.737 -57.343  1.00  5.97 C
ATOM    785  CD2 LEU   104A    -37.265  83.977 -59.039  1.00  1.35 C
ATOM    786  C   LEU   104A    -33.050  85.371 -58.872  1.00  4.91 C
ATOM    787  O   LEU   104A    -32.307  86.045 -58.156  1.00  7.01 O
ATOM    788  N   LYS   105A    -32.896  85.285 -60.184  1.00  5.26 N
ATOM    789  CA  LYS   105A    -31.843  86.024 -60.863  1.00  8.68 C
ATOM    790  CB  LYS   105A    -30.539  85.222 -60.885  1.00  9.23 C
```

FIGURE 9 - 13

```
ATOM    791  CG  LYS   105A    -29.405  85.974 -61.565  1.00 14.33 C
ATOM    792  CD  LYS   105A    -28.132  85.159 -61.638  1.00 18.70 C
ATOM    793  CE  LYS   105A    -27.089  85.880 -62.481  1.00 19.07 C
ATOM    794  NZ  LYS   105A    -25.876  85.052 -62.695  1.00 21.33 N
ATOM    795  C   LYS   105A    -32.264  86.337 -62.284  1.00  8.50 C
ATOM    796  O   LYS   105A    -32.954  85.551 -62.923  1.00  7.17 O
ATOM    797  N   THR   106A    -31.850  87.494 -62.776  1.00  9.20 N
ATOM    798  CA  THR   106A    -32.187  87.881 -64.134  1.00 11.88 C
ATOM    799  CB  THR   106A    -32.861  89.271 -64.162  1.00 13.82 C
ATOM    800  OG1 THR   106A    -33.007  89.710 -65.517  1.00 16.55 O
ATOM    801  CG2 THR   106A    -32.036  90.275 -63.393  1.00 15.81 C
ATOM    802  C   THR   106A    -30.913  87.892 -64.972  1.00 12.59 C
ATOM    803  O   THR   106A    -29.881  88.418 -64.547  1.00  9.88 O
ATOM    804  N   GLY   107A    -30.984  87.280 -66.148  1.00 12.58 N
ATOM    805  CA  GLY   107A    -29.832  87.234 -67.026  1.00 15.66 C
ATOM    806  C   GLY   107A    -28.661  86.457 -66.461  1.00 16.09 C
ATOM    807  O   GLY   107A    -28.804  85.671 -65.522  1.00 16.17 O
ATOM    808  N   GLY   108A    -27.488  86.682 -67.038  1.00 18.04 N
ATOM    809  CA  GLY   108A    -26.302  85.986 -66.583  1.00 17.36 C
ATOM    810  C   GLY   108A    -26.044  84.761 -67.438  1.00 18.87 C
ATOM    811  O   GLY   108A    -26.798  84.479 -68.375  1.00 18.34 O
ATOM    812  N   GLY   109A    -24.986  84.028 -67.107  1.00 18.26 N
ATOM    813  CA  GLY   109A    -24.640  82.837 -67.862  1.00 17.43 C
ATOM    814  C   GLY   109A    -25.479  81.620 -67.518  1.00 16.37 C
ATOM    815  O   GLY   109A    -26.601  81.738 -67.028  1.00 16.42 O
ATOM    816  N   HIS   110A    -24.928  80.441 -67.775  1.00 15.39 N
ATOM    817  CA  HIS   110A    -25.626  79.193 -67.500  1.00 15.16 C
ATOM    818  CB  HIS   110A    -25.635  78.318 -68.754  1.00 14.88 C
ATOM    819  CG  HIS   110A    -24.282  78.129 -69.362  1.00 15.15 C
ATOM    820  CD2 HIS   110A    -23.803  78.442 -70.589  1.00 14.18 C
ATOM    821  ND1 HIS   110A    -23.234  77.555 -68.677  1.00 14.24 N
ATOM    822  CE1 HIS   110A    -22.167  77.524 -69.456  1.00 15.94 C
ATOM    823  NE2 HIS   110A    -22.486  78.056 -70.621  1.00 13.76 N
ATOM    824  C   HIS   110A    -24.989  78.429 -66.345  1.00 15.27 C
ATOM    825  O   HIS   110A    -25.562  77.466 -65.841  1.00 13.41 O
ATOM    826  N   GLY   111A    -23.803  78.865 -65.927  1.00 16.21 N
ATOM    827  CA  GLY   111A    -23.113  78.207 -64.830  1.00 15.78 C
ATOM    828  C   GLY   111A    -22.869  76.726 -65.053  1.00 16.47 C
ATOM    829  O   GLY   111A    -22.812  75.950 -64.100  1.00 16.61 O
ATOM    830  N   GLY   112A    -22.717  76.329 -66.314  1.00 15.55 N
ATOM    831  CA  GLY   112A    -22.483  74.931 -66.623  1.00 15.17 C
ATOM    832  C   GLY   112A    -23.753  74.107 -66.755  1.00 13.15 C
ATOM    833  O   GLY   112A    -23.702  72.948 -67.166  1.00 16.31 O
ATOM    834  N   HIS   113A    -24.896  74.692 -66.409  1.00 12.94 N
ATOM    835  CA  HIS   113A    -26.158  73.973 -66.512  1.00 10.81 C
ATOM    836  CB  HIS   113A    -27.266  74.722 -65.770  1.00  8.88 C
ATOM    837  CG  HIS   113A    -28.505  73.907 -65.571  1.00  6.89 C
ATOM    838  CD2 HIS   113A    -29.075  73.409 -64.448  1.00  5.22 C
ATOM    839  ND1 HIS   113A    -29.289  73.480 -66.620  1.00  8.12 N
ATOM    840  CE1 HIS   113A    -30.287  72.752 -66.154  1.00  5.50 C
ATOM    841  NE2 HIS   113A    -30.181  72.692 -64.837  1.00  5.53 N
ATOM    842  C   HIS   113A    -26.491  73.858 -67.995  1.00  9.87 C
ATOM    843  O   HIS   113A    -26.792  74.850 -68.656  1.00 10.05 O
ATOM    844  N   ASN   114A    -26.422  72.638 -68.512  1.00 10.16 N
ATOM    845  CA  ASN   114A    -26.667  72.392 -69.923  1.00 10.50 C
ATOM    846  CB  ASN   114A    -26.256  70.959 -70.251  1.00 14.98 C
ATOM    847  CG  ASN   114A    -24.774  70.736 -70.026  1.00 18.32 C
ATOM    848  OD1 ASN   114A    -24.314  69.608 -69.857  1.00 20.80 O
ATOM    849  ND2 ASN   114A    -24.014  71.828 -70.024  1.00 19.91 N
ATOM    850  C   ASN   114A    -28.078  72.692 -70.397  1.00  9.08 C
ATOM    851  O   ASN   114A    -28.305  72.873 -71.593  1.00  6.27 O
```

FIGURE 9 - 14

```
ATOM    852  N   GLY   115A    -29.021  72.755 -69.461  1.00  7.09 N
ATOM    853  CA  GLY   115A    -30.384  73.081 -69.833  1.00  6.38 C
ATOM    854  C   GLY   115A    -30.415  74.569 -70.121  1.00  4.56 C
ATOM    855  O   GLY   115A    -30.936  75.016 -71.147  1.00  3.76 O
ATOM    856  N   LEU   116A    -29.842  75.345 -69.207  1.00  5.41 N
ATOM    857  CA  LEU   116A    -29.785  76.789 -69.362  1.00  5.93 C
ATOM    858  CB  LEU   116A    -29.159  77.444 -68.122  1.00  7.99 C
ATOM    859  CG  LEU   116A    -29.963  77.454 -66.819  1.00 11.76 C
ATOM    860  CD1 LEU   116A    -29.132  78.075 -65.701  1.00 12.07 C
ATOM    861  CD2 LEU   116A    -31.254  78.249 -67.025  1.00 11.88 C
ATOM    862  C   LEU   116A    -28.946  77.119 -70.593  1.00  5.55 C
ATOM    863  O   LEU   116A    -29.282  78.014 -71.360  1.00  4.53 O
ATOM    864  N   LYS   117A    -27.859  76.377 -70.785  1.00  5.46 N
ATOM    865  CA  LYS   117A    -26.984  76.609 -71.928  1.00  6.12 C
ATOM    866  CB  LYS   117A    -25.824  75.601 -71.927  1.00  5.51 C
ATOM    867  CG  LYS   117A    -24.880  75.759 -73.114  1.00  8.75 C
ATOM    868  CD  LYS   117A    -23.552  75.025 -72.915  1.00  4.61 C
ATOM    869  CE  LYS   117A    -23.698  73.527 -72.753  1.00  9.26 C
ATOM    870  NZ  LYS   117A    -22.348  72.894 -72.555  1.00 11.81 N
ATOM    871  C   LYS   117A    -27.735  76.530 -73.261  1.00  6.70 C
ATOM    872  O   LYS   117A    -27.581  77.405 -74.117  1.00  7.74 O
ATOM    873  N   ASP   118A    -28.554  75.495 -73.426  1.00  6.22 N
ATOM    874  CA  ASP   118A    -29.304  75.304 -74.665  1.00  5.45 C
ATOM    875  CB  ASP   118A    -29.932  73.902 -74.693  1.00  2.99 C
ATOM    876  CG  ASP   118A    -30.352  73.481 -76.092  1.00  3.70 C
ATOM    877  OD1 ASP   118A    -29.478  73.426 -76.982  1.00  2.25 O
ATOM    878  OD2 ASP   118A    -31.549  73.206 -76.312  1.00  1.28 O
ATOM    879  C   ASP   118A    -30.387  76.366 -74.836  1.00  4.76 C
ATOM    880  O   ASP   118A    -30.643  76.841 -75.943  1.00  5.67 O
ATOM    881  N   ILE   119A    -31.035  76.732 -73.739  1.00  4.75 N
ATOM    882  CA  ILE   119A    -32.070  77.745 -73.807  1.00  4.40 C
ATOM    883  CB  ILE   119A    -32.754  77.936 -72.441  1.00  2.72 C
ATOM    884  CG2 ILE   119A    -33.725  79.105 -72.506  1.00  2.29 C
ATOM    885  CG1 ILE   119A    -33.510  76.653 -72.069  1.00  2.63 C
ATOM    886  CD1 ILE   119A    -34.062  76.649 -70.655  1.00  2.60 C
ATOM    887  C   ILE   119A    -31.454  79.059 -74.272  1.00  5.25 C
ATOM    888  O   ILE   119A    -31.993  79.725 -75.156  1.00  5.20 O
ATOM    889  N   ILE   120A    -30.323  79.428 -73.682  1.00  6.96 N
ATOM    890  CA  ILE   120A    -29.660  80.665 -74.078  1.00  9.05 C
ATOM    891  CB  ILE   120A    -28.400  80.937 -73.238  1.00  9.93 C
ATOM    892  CG2 ILE   120A    -27.643  82.129 -73.810  1.00 12.52 C
ATOM    893  CG1 ILE   120A    -28.789  81.217 -71.787  1.00 12.36 C
ATOM    894  CD1 ILE   120A    -27.605  81.385 -70.862  1.00 12.03 C
ATOM    895  C   ILE   120A    -29.257  80.589 -75.546  1.00  9.38 C
ATOM    896  O   ILE   120A    -29.477  81.531 -76.301  1.00 12.07 O
ATOM    897  N   ALA   121A    -28.676  79.464 -75.950  1.00  9.92 N
ATOM    898  CA  ALA   121A    -28.240  79.293 -77.334  1.00  9.79 C
ATOM    899  CB  ALA   121A    -27.520  77.959 -77.501  1.00  9.99 C
ATOM    900  C   ALA   121A    -29.415  79.372 -78.298  1.00 12.25 C
ATOM    901  O   ALA   121A    -29.327  79.995 -79.356  1.00 13.32 O
ATOM    902  N   GLN   122A    -30.522  78.746 -77.922  1.00 12.36 N
ATOM    903  CA  GLN   122A    -31.700  78.734 -78.770  1.00 17.39 C
ATOM    904  CB  GLN   122A    -32.719  77.735 -78.224  1.00 20.81 C
ATOM    905  CG  GLN   122A    -33.650  77.198 -79.283  1.00 26.52 C
ATOM    906  CD  GLN   122A    -32.900  76.490 -80.388  1.00 28.82 C
ATOM    907  OE1 GLN   122A    -32.330  75.417 -80.181  1.00 31.07 O
ATOM    908  NE2 GLN   122A    -32.884  77.094 -81.571  1.00 31.63 N
ATOM    909  C   GLN   122A    -32.348  80.107 -78.916  1.00 18.24 C
ATOM    910  O   GLN   122A    -32.781  80.480 -80.004  1.00 18.95 O
ATOM    911  N   LEU   123A    -32.416  80.865 -77.827  1.00 17.48 N
ATOM    912  CA  LEU   123A    -33.036  82.181 -77.893  1.00 20.57 C
```

FIGURE 9 - 15

```
ATOM    913  CB  LEU  123A    -33.470  82.635 -76.494  1.00 15.46 C
ATOM    914  CG  LEU  123A    -34.530  81.734 -75.841  1.00 15.80 C
ATOM    915  CD1 LEU  123A    -34.938  82.302 -74.494  1.00 12.14 C
ATOM    916  CD2 LEU  123A    -35.740  81.609 -76.759  1.00 14.23 C
ATOM    917  C   LEU  123A    -32.096  83.200 -78.532  1.00 24.33 C
ATOM    918  O   LEU  123A    -32.019  84.352 -78.102  1.00 24.22 O
ATOM    919  N   GLY  124A    -31.393  82.760 -79.573  1.00 27.34 N
ATOM    920  CA  GLY  124A    -30.464  83.625 -80.278  1.00 31.47 C
ATOM    921  C   GLY  124A    -29.503  84.337 -79.349  1.00 33.87 C
ATOM    922  O   GLY  124A    -29.284  85.542 -79.479  1.00 36.58 O
ATOM    923  N   ASN  125A    -28.929  83.591 -78.410  1.00 35.87 N
ATOM    924  CA  ASN  125A    -27.989  84.147 -77.443  1.00 35.97 C
ATOM    925  CB  ASN  125A    -26.801  84.796 -78.163  1.00 38.93 C
ATOM    926  CG  ASN  125A    -25.827  83.774 -78.730  1.00 40.22 C
ATOM    927  OD1 ASN  125A    -24.886  84.128 -79.443  1.00 41.68 O
ATOM    928  ND2 ASN  125A    -26.043  82.503 -78.407  1.00 40.73 N
ATOM    929  C   ASN  125A    -28.650  85.167 -76.520  1.00 34.54 C
ATOM    930  O   ASN  125A    -28.031  86.160 -76.135  1.00 35.18 O
ATOM    931  N   ASN  126A    -29.909  84.921 -76.170  1.00 33.49 N
ATOM    932  CA  ASN  126A    -30.638  85.817 -75.280  1.00 33.84 C
ATOM    933  CB  ASN  126A    -32.002  86.183 -75.872  1.00 32.37 C
ATOM    934  CG  ASN  126A    -31.907  87.262 -76.933  1.00 31.11 C
ATOM    935  OD1 ASN  126A    -32.918  87.704 -77.476  1.00 30.83 O
ATOM    936  ND2 ASN  126A    -30.686  87.694 -77.231  1.00 32.10 N
ATOM    937  C   ASN  126A    -30.831  85.210 -73.898  1.00 34.67 C
ATOM    938  O   ASN  126A    -31.801  84.496 -73.650  1.00 33.44 O
ATOM    939  N   ASN  127A    -29.897  85.502 -73.000  1.00 34.58 N
ATOM    940  CA  ASN  127A    -29.962  84.995 -71.635  1.00 34.82 C
ATOM    941  CB  ASN  127A    -28.550  84.797 -71.081  1.00 39.11 C
ATOM    942  CG  ASN  127A    -27.732  86.074 -71.103  1.00 42.27 C
ATOM    943  OD1 ASN  127A    -27.482  86.646 -72.164  1.00 43.93 O
ATOM    944  ND2 ASN  127A    -27.311  86.528 -69.927  1.00 45.03 N
ATOM    945  C   ASN  127A    -30.723  85.992 -70.769  1.00 31.48 C
ATOM    946  O   ASN  127A    -30.757  85.872 -69.543  1.00 32.55 O
ATOM    947  N   SER  128A    -31.339  86.971 -71.426  1.00 27.79 N
ATOM    948  CA  SER  128A    -32.094  88.014 -70.745  1.00 22.25 C
ATOM    949  CB  SER  128A    -32.316  89.194 -71.696  1.00 23.86 C
ATOM    950  OG  SER  128A    -32.902  88.768 -72.914  1.00 27.87 O
ATOM    951  C   SER  128A    -33.433  87.545 -70.167  1.00 17.21 C
ATOM    952  O   SER  128A    -34.367  88.334 -70.032  1.00 16.09 O
ATOM    953  N   PHE  129A    -33.525  86.263 -69.828  1.00 11.41 N
ATOM    954  CA  PHE  129A    -34.744  85.729 -69.229  1.00  9.15 C
ATOM    955  CB  PHE  129A    -35.115  84.383 -69.857  1.00  5.10 C
ATOM    956  CG  PHE  129A    -33.998  83.385 -69.854  1.00  3.64 C
ATOM    957  CD1 PHE  129A    -33.822  82.517 -68.781  1.00  3.05 C
ATOM    958  CD2 PHE  129A    -33.124  83.306 -70.932  1.00  3.02 C
ATOM    959  CE1 PHE  129A    -32.792  81.578 -68.785  1.00  6.64 C
ATOM    960  CE2 PHE  129A    -32.089  82.372 -70.948  1.00  4.68 C
ATOM    961  CZ  PHE  129A    -31.923  81.507 -69.875  1.00  6.52 C
ATOM    962  C   PHE  129A    -34.471  85.568 -67.739  1.00  5.65 C
ATOM    963  O   PHE  129A    -33.322  85.598 -67.317  1.00  6.79 O
ATOM    964  N   HIS  130A    -35.521  85.393 -66.943  1.00  5.39 N
ATOM    965  CA  HIS  130A    -35.352  85.262 -65.505  1.00  3.48 C
ATOM    966  CB  HIS  130A    -36.510  85.959 -64.796  1.00  3.77 C
ATOM    967  CG  HIS  130A    -36.729  87.359 -65.274  1.00  6.65 C
ATOM    968  CD2 HIS  130A    -36.193  88.531 -64.864  1.00  6.30 C
ATOM    969  ND1 HIS  130A    -37.525  87.659 -66.358  1.00  8.43 N
ATOM    970  CE1 HIS  130A    -37.469  88.957 -66.597  1.00  5.55 C
ATOM    971  NE2 HIS  130A    -36.667  89.509 -65.706  1.00  9.03 N
ATOM    972  C   HIS  130A    -35.231  83.814 -65.063  1.00  3.06 C
ATOM    973  O   HIS  130A    -35.746  82.909 -65.722  1.00  1.63 O
```

FIGURE 9 - 16

```
ATOM    974  N   ARG  131A    -34.537  83.596 -63.951  1.00  3.89 N
ATOM    975  CA  ARG  131A    -34.340  82.247 -63.448  1.00  3.99 C
ATOM    976  CB  ARG  131A    -32.882  81.808 -63.650  1.00  5.92 C
ATOM    977  CG  ARG  131A    -32.337  81.973 -65.064  1.00  7.19 C
ATOM    978  CD  ARG  131A    -31.906  83.401 -65.336  1.00 10.69 C
ATOM    979  NE  ARG  131A    -31.195  83.549 -66.606  1.00 10.32 N
ATOM    980  CZ  ARG  131A    -29.961  83.110 -66.842  1.00 12.78 C
ATOM    981  NH1 ARG  131A    -29.277  82.480 -65.895  1.00 10.23 N
ATOM    982  NH2 ARG  131A    -29.394  83.326 -68.025  1.00 13.36 N
ATOM    983  C   ARG  131A    -34.691  82.067 -61.975  1.00  4.09 C
ATOM    984  O   ARG  131A    -34.407  82.933 -61.144  1.00  3.82 O
ATOM    985  N   LEU  132A    -35.318  80.935 -61.667  1.00  3.73 N
ATOM    986  CA  LEU  132A    -35.665  80.591 -60.296  1.00  3.43 C
ATOM    987  CB  LEU  132A    -37.093  80.027 -60.216  1.00  5.42 C
ATOM    988  CG  LEU  132A    -37.792  80.073 -58.848  1.00  9.25 C
ATOM    989  CD1 LEU  132A    -39.131  79.330 -58.902  1.00  8.43 C
ATOM    990  CD2 LEU  132A    -36.917  79.449 -57.805  1.00 10.45 C
ATOM    991  C   LEU  132A    -34.635  79.504 -59.973  1.00  3.69 C
ATOM    992  O   LEU  132A    -34.710  78.389 -60.494  1.00  1.00 O
ATOM    993  N   ARG  133A    -33.657  79.841 -59.141  1.00  1.94 N
ATOM    994  CA  ARG  133A    -32.596  78.902 -58.788  1.00  2.71 C
ATOM    995  CB  ARG  133A    -31.289  79.664 -58.572  1.00  3.22 C
ATOM    996  CG  ARG  133A    -30.960  80.589 -59.729  1.00  5.14 C
ATOM    997  CD  ARG  133A    -29.811  81.513 -59.404  1.00  7.26 C
ATOM    998  NE  ARG  133A    -28.545  80.802 -59.292  1.00  7.21 N
ATOM    999  CZ  ARG  133A    -27.380  81.397 -59.075  1.00 10.51 C
ATOM   1000  NH1 ARG  133A    -27.329  82.715 -58.949  1.00 10.82 N
ATOM   1001  NH2 ARG  133A    -26.266  80.678 -58.983  1.00 12.00 N
ATOM   1002  C   ARG  133A    -32.944  78.104 -57.545  1.00  4.29 C
ATOM   1003  O   ARG  133A    -33.007  78.653 -56.448  1.00  3.68 O
ATOM   1004  N   LEU  134A    -33.154  76.805 -57.739  1.00  1.00 N
ATOM   1005  CA  LEU  134A    -33.516  75.891 -56.662  1.00  2.72 C
ATOM   1006  CB  LEU  134A    -34.602  74.934 -57.155  1.00  4.09 C
ATOM   1007  CG  LEU  134A    -35.825  75.658 -57.726  1.00  4.01 C
ATOM   1008  CD1 LEU  134A    -36.759  74.644 -58.370  1.00  1.00 C
ATOM   1009  CD2 LEU  134A    -36.537  76.430 -56.615  1.00  6.05 C
ATOM   1010  C   LEU  134A    -32.308  75.106 -56.167  1.00  2.50 C
ATOM   1011  O   LEU  134A    -31.837  74.181 -56.832  1.00  2.89 O
ATOM   1012  N   GLY  135A    -31.818  75.477 -54.990  1.00  3.27 N
ATOM   1013  CA  GLY  135A    -30.653  74.810 -54.430  1.00  2.97 C
ATOM   1014  C   GLY  135A    -30.758  73.303 -54.287  1.00  1.55 C
ATOM   1015  O   GLY  135A    -31.782  72.779 -53.853  1.00  2.43 O
ATOM   1016  N   ILE  136A    -29.688  72.600 -54.652  1.00  1.80 N
ATOM   1017  CA  ILE  136A    -29.647  71.144 -54.546  1.00  2.11 C
ATOM   1018  CB  ILE  136A    -29.766  70.457 -55.924  1.00  3.61 C
ATOM   1019  CG2 ILE  136A    -31.155  70.681 -56.503  1.00  3.23 C
ATOM   1020  CG1 ILE  136A    -28.692  70.992 -56.874  1.00  1.00 C
ATOM   1021  CD1 ILE  136A    -28.394  70.046 -58.019  1.00  1.30 C
ATOM   1022  C   ILE  136A    -28.317  70.714 -53.933  1.00  1.47 C
ATOM   1023  O   ILE  136A    -28.098  69.539 -53.655  1.00  1.00 O
ATOM   1024  N   GLY  137A    -27.430  71.680 -53.722  1.00  2.06 N
ATOM   1025  CA  GLY  137A    -26.127  71.369 -53.158  1.00  1.97 C
ATOM   1026  C   GLY  137A    -25.098  71.000 -54.218  1.00  2.24 C
ATOM   1027  O   GLY  137A    -25.441  70.800 -55.384  1.00  1.00 O
ATOM   1028  N   HIS  138A    -23.833  70.931 -53.809  1.00  1.38 N
ATOM   1029  CA  HIS  138A    -22.733  70.581 -54.709  1.00  2.60 C
ATOM   1030  CB  HIS  138A    -21.778  71.766 -54.867  1.00  3.56 C
ATOM   1031  CG  HIS  138A    -20.804  71.607 -55.993  1.00  5.13 C
ATOM   1032  CD2 HIS  138A    -19.728  70.796 -56.131  1.00  4.96 C
ATOM   1033  ND1 HIS  138A    -20.910  72.306 -57.177  1.00  6.28 N
ATOM   1034  CE1 HIS  138A    -19.942  71.932 -57.995  1.00  4.76 C
```

FIGURE 9-17

```
ATOM   1035  NE2 HIS   138A    -19.212  71.015 -57.384  1.00  7.39 N
ATOM   1036  C   HIS   138A    -21.976  69.383 -54.113  1.00  3.77 C
ATOM   1037  O   HIS   138A    -21.717  69.347 -52.912  1.00  5.10 O
ATOM   1038  N   PRO   139A    -21.614  68.387 -54.945  1.00  2.94 N
ATOM   1039  CD  PRO   139A    -21.846  68.295 -56.399  1.00  3.47 C
ATOM   1040  CA  PRO   139A    -20.894  67.202 -54.465  1.00  3.90 C
ATOM   1041  CB  PRO   139A    -20.976  66.250 -55.654  1.00  2.60 C
ATOM   1042  CG  PRO   139A    -20.880  67.188 -56.811  1.00  2.19 C
ATOM   1043  C   PRO   139A    -19.452  67.473 -54.056  1.00  5.14 C
ATOM   1044  O   PRO   139A    -18.816  66.632 -53.428  1.00  4.10 O
ATOM   1045  N   GLY   140A    -18.945  68.643 -54.428  1.00  6.91 N
ATOM   1046  CA  GLY   140A    -17.583  69.013 -54.102  1.00 11.14 C
ATOM   1047  C   GLY   140A    -16.783  69.298 -55.361  1.00 14.48 C
ATOM   1048  O   GLY   140A    -16.739  70.434 -55.843  1.00 16.27 O
ATOM   1049  N   HIS   141A    -16.160  68.264 -55.913  1.00 13.75 N
ATOM   1050  CA  HIS   141A    -15.356  68.426 -57.116  1.00 13.95 C
ATOM   1051  CB  HIS   141A    -14.313  67.313 -57.209  1.00 15.58 C
ATOM   1052  CG  HIS   141A    -13.281  67.552 -58.266  1.00 17.24 C
ATOM   1053  CD2 HIS   141A    -13.082  66.962 -59.467  1.00 16.54 C
ATOM   1054  ND1 HIS   141A    -12.320  68.535 -58.158  1.00 17.49 N
ATOM   1055  CE1 HIS   141A    -11.575  68.539 -59.248  1.00 16.38 C
ATOM   1056  NE2 HIS   141A    -12.017  67.595 -60.059  1.00 18.18 N
ATOM   1057  C   HIS   141A    -16.219  68.421 -58.374  1.00 13.82 C
ATOM   1058  O   HIS   141A    -17.232  67.723 -58.441  1.00 14.66 O
ATOM   1059  N   SER   142A    -15.805  69.200 -59.370  1.00 12.96 N
ATOM   1060  CA  SER   142A    -16.534  69.310 -60.631  1.00 11.02 C
ATOM   1061  CB  SER   142A    -15.802  70.268 -61.579  1.00  9.93 C
ATOM   1062  OG  SER   142A    -14.481  69.823 -61.849  1.00  9.22 O
ATOM   1063  C   SER   142A    -16.756  67.975 -61.333  1.00 12.17 C
ATOM   1064  O   SER   142A    -17.763  67.790 -62.022  1.00  9.42 O
ATOM   1065  N   SER   143A    -15.823  67.045 -61.155  1.00 11.55 N
ATOM   1066  CA  SER   143A    -15.926  65.735 -61.793  1.00 11.74 C
ATOM   1067  CB  SER   143A    -14.622  64.948 -61.609  1.00 12.13 C
ATOM   1068  OG  SER   143A    -14.411  64.622 -60.246  1.00 12.40 O
ATOM   1069  C   SER   143A    -17.089  64.886 -61.284  1.00 11.16 C
ATOM   1070  O   SER   143A    -17.521  63.954 -61.967  1.00 10.69 O
ATOM   1071  N   LEU   144A    -17.591  65.203 -60.091  1.00  9.26 N
ATOM   1072  CA  LEU   144A    -18.691  64.443 -59.498  1.00  7.48 C
ATOM   1073  CB  LEU   144A    -18.505  64.343 -57.983  1.00  7.45 C
ATOM   1074  CG  LEU   144A    -17.321  63.550 -57.421  1.00  8.05 C
ATOM   1075  CD1 LEU   144A    -17.299  63.714 -55.905  1.00  9.06 C
ATOM   1076  CD2 LEU   144A    -17.456  62.075 -57.787  1.00  8.30 C
ATOM   1077  C   LEU   144A    -20.077  65.023 -59.774  1.00  7.06 C
ATOM   1078  O   LEU   144A    -21.090  64.434 -59.388  1.00  4.53 O
ATOM   1079  N   VAL   145A    -20.119  66.163 -60.449  1.00  6.61 N
ATOM   1080  CA  VAL   145A    -21.378  66.830 -60.739  1.00  7.34 C
ATOM   1081  CB  VAL   145A    -21.124  68.174 -61.464  1.00  7.28 C
ATOM   1082  CG1 VAL   145A    -22.442  68.853 -61.792  1.00  8.03 C
ATOM   1083  CG2 VAL   145A    -20.278  69.084 -60.580  1.00 10.51 C
ATOM   1084  C   VAL   145A    -22.394  66.005 -61.532  1.00  8.25 C
ATOM   1085  O   VAL   145A    -23.554  65.916 -61.134  1.00  7.56 O
ATOM   1086  N   SER   146A    -21.977  65.406 -62.643  1.00  6.30 N
ATOM   1087  CA  SER   146A    -22.898  64.607 -63.456  1.00  9.28 C
ATOM   1088  CB  SER   146A    -22.155  63.942 -64.619  1.00 13.00 C
ATOM   1089  OG  SER   146A    -21.782  64.894 -65.600  1.00 14.89 O
ATOM   1090  C   SER   146A    -23.643  63.536 -62.664  1.00  8.27 C
ATOM   1091  O   SER   146A    -24.873  63.477 -62.701  1.00  9.47 O
ATOM   1092  N   GLY   147A    -22.906  62.688 -61.955  1.00  5.67 N
ATOM   1093  CA  GLY   147A    -23.545  61.641 -61.181  1.00  5.64 C
ATOM   1094  C   GLY   147A    -24.359  62.185 -60.021  1.00  5.45 C
ATOM   1095  O   GLY   147A    -25.291  61.545 -59.540  1.00  4.99 O
```

FIGURE 9 - 18

```
ATOM   1096  N   TYR  148A   -24.009  63.378 -59.566  1.00  4.05 N
ATOM   1097  CA  TYR  148A   -24.712  63.989 -58.448  1.00  4.39 C
ATOM   1098  CB  TYR  148A   -23.875  65.136 -57.888  1.00  3.04 C
ATOM   1099  CG  TYR  148A   -24.486  65.820 -56.685  1.00  4.05 C
ATOM   1100  CD1 TYR  148A   -24.334  65.290 -55.401  1.00  2.82 C
ATOM   1101  CE1 TYR  148A   -24.871  65.934 -54.293  1.00  3.66 C
ATOM   1102  CD2 TYR  148A   -25.196  67.010 -56.830  1.00  1.00 C
ATOM   1103  CE2 TYR  148A   -25.737  67.662 -55.733  1.00  2.06 C
ATOM   1104  CZ  TYR  148A   -25.571  67.121 -54.470  1.00  3.84 C
ATOM   1105  OH  TYR  148A   -26.101  67.771 -53.390  1.00  2.60 O
ATOM   1106  C   TYR  148A   -26.098  64.507 -58.829  1.00  4.46 C
ATOM   1107  O   TYR  148A   -27.084  64.218 -58.147  1.00  4.47 O
ATOM   1108  N   VAL  149A   -26.177  65.269 -59.915  1.00  3.43 N
ATOM   1109  CA  VAL  149A   -27.460  65.827 -60.329  1.00  1.00 C
ATOM   1110  CB  VAL  149A   -27.291  66.990 -61.342  1.00  2.54 C
ATOM   1111  CG1 VAL  149A   -26.455  68.098 -60.718  1.00  2.54 C
ATOM   1112  CG2 VAL  149A   -26.642  66.485 -62.625  1.00  3.20 C
ATOM   1113  C   VAL  149A   -28.355  64.767 -60.938  1.00  2.51 C
ATOM   1114  O   VAL  149A   -29.562  64.948 -61.023  1.00  1.00 O
ATOM   1115  N   LEU  150A   -27.763  63.655 -61.351  1.00  1.00 N
ATOM   1116  CA  LEU  150A   -28.547  62.579 -61.941  1.00  3.02 C
ATOM   1117  CB  LEU  150A   -27.760  61.902 -63.062  1.00  3.49 C
ATOM   1118  CG  LEU  150A   -27.362  62.841 -64.206  1.00  5.59 C
ATOM   1119  CD1 LEU  150A   -26.635  62.051 -65.276  1.00  6.87 C
ATOM   1120  CD2 LEU  150A   -28.597  63.535 -64.777  1.00  4.33 C
ATOM   1121  C   LEU  150A   -28.928  61.566 -60.874  1.00  3.21 C
ATOM   1122  O   LEU  150A   -29.521  60.530 -61.167  1.00  1.39 O
ATOM   1123  N   GLY  151A   -28.582  61.871 -59.628  1.00  3.02 N
ATOM   1124  CA  GLY  151A   -28.910  60.966 -58.545  1.00  5.57 C
ATOM   1125  C   GLY  151A   -30.136  61.411 -57.777  1.00  6.11 C
ATOM   1126  O   GLY  151A   -30.537  62.568 -57.852  1.00  3.26 O
ATOM   1127  N   ARG  152A   -30.752  60.487 -57.052  1.00  7.64 N
ATOM   1128  CA  ARG  152A   -31.919  60.827 -56.256  1.00  9.22 C
ATOM   1129  CB  ARG  152A   -32.901  59.654 -56.208  1.00 13.25 C
ATOM   1130  CG  ARG  152A   -33.505  59.338 -57.567  1.00 17.92 C
ATOM   1131  CD  ARG  152A   -34.493  58.183 -57.505  1.00 22.67 C
ATOM   1132  NE  ARG  152A   -34.878  57.717 -58.837  1.00 27.01 N
ATOM   1133  CZ  ARG  152A   -35.584  58.426 -59.714  1.00 28.52 C
ATOM   1134  NH1 ARG  152A   -35.995  59.648 -59.411  1.00 28.94 N
ATOM   1135  NH2 ARG  152A   -35.888  57.911 -60.897  1.00 28.72 N
ATOM   1136  C   ARG  152A   -31.421  61.183 -54.867  1.00  7.91 C
ATOM   1137  O   ARG  152A   -30.754  60.391 -54.206  1.00  9.42 O
ATOM   1138  N   ALA  153A   -31.738  62.395 -54.433  1.00  7.52 N
ATOM   1139  CA  ALA  153A   -31.289  62.868 -53.136  1.00  9.09 C
ATOM   1140  CB  ALA  153A   -31.548  64.363 -53.024  1.00  8.87 C
ATOM   1141  C   ALA  153A   -31.956  62.147 -51.975  1.00  8.66 C
ATOM   1142  O   ALA  153A   -33.076  61.646 -52.102  1.00  9.81 O
ATOM   1143  N   PRO  154A   -31.259  62.063 -50.828  1.00 10.19 N
ATOM   1144  CD  PRO  154A   -29.884  62.512 -50.555  1.00 10.91 C
ATOM   1145  CA  PRO  154A   -31.839  61.398 -49.657  1.00 10.87 C
ATOM   1146  CB  PRO  154A   -30.704  61.445 -48.629  1.00 10.36 C
ATOM   1147  CG  PRO  154A   -29.882  62.622 -49.053  1.00 12.30 C
ATOM   1148  C   PRO  154A   -33.060  62.226 -49.255  1.00 10.57 C
ATOM   1149  O   PRO  154A   -33.146  63.404 -49.608  1.00  7.67 O
ATOM   1150  N   ARG  155A   -34.000  61.633 -48.528  1.00 11.24 N
ATOM   1151  CA  ARG  155A   -35.212  62.358 -48.161  1.00 15.17 C
ATOM   1152  CB  ARG  155A   -36.145  61.470 -47.340  1.00 19.61 C
ATOM   1153  CG  ARG  155A   -35.729  61.300 -45.899  1.00 26.57 C
ATOM   1154  CD  ARG  155A   -36.931  60.958 -45.052  1.00 32.23 C
ATOM   1155  NE  ARG  155A   -36.618  60.967 -43.629  1.00 37.23 N
ATOM   1156  CZ  ARG  155A   -36.168  62.030 -42.968  1.00 39.78 C
```

FIGURE 9 - 19

ATOM	1157	NH1	ARG	155A	-35.970	63.183	-43.600	1.00	39.27	N
ATOM	1158	NH2	ARG	155A	-35.920	61.941	-41.668	1.00	41.07	N
ATOM	1159	C	ARG	155A	-34.996	63.677	-47.422	1.00	11.94	C
ATOM	1160	O	ARG	155A	-35.750	64.625	-47.623	1.00	11.79	O
ATOM	1161	N	SER	156A	-33.978	63.749	-46.572	1.00	11.69	N
ATOM	1162	CA	SER	156A	-33.726	64.984	-45.835	1.00	10.21	C
ATOM	1163	CB	SER	156A	-32.522	64.821	-44.903	1.00	13.03	C
ATOM	1164	OG	SER	156A	-31.327	64.615	-45.634	1.00	14.20	O
ATOM	1165	C	SER	156A	-33.483	66.153	-46.789	1.00	9.66	C
ATOM	1166	O	SER	156A	-33.881	67.283	-46.512	1.00	8.09	O
ATOM	1167	N	GLU	157A	-32.833	65.876	-47.915	1.00	7.66	N
ATOM	1168	CA	GLU	157A	-32.550	66.910	-48.904	1.00	6.24	C
ATOM	1169	CB	GLU	157A	-31.357	66.496	-49.767	1.00	4.69	C
ATOM	1170	CG	GLU	157A	-30.028	66.524	-49.030	1.00	6.65	C
ATOM	1171	CD	GLU	157A	-28.879	66.013	-49.881	1.00	9.21	C
ATOM	1172	OE1	GLU	157A	-28.948	66.147	-51.122	1.00	8.99	O
ATOM	1173	OE2	GLU	157A	-27.896	65.491	-49.312	1.00	6.68	O
ATOM	1174	C	GLU	157A	-33.773	67.161	-49.783	1.00	6.40	C
ATOM	1175	O	GLU	157A	-34.040	68.296	-50.193	1.00	4.25	O
ATOM	1176	N	GLN	158A	-34.515	66.097	-50.075	1.00	5.53	N
ATOM	1177	CA	GLN	158A	-35.717	66.226	-50.885	1.00	6.92	C
ATOM	1178	CB	GLN	158A	-36.370	64.859	-51.093	1.00	7.82	C
ATOM	1179	CG	GLN	158A	-37.654	64.918	-51.896	1.00	12.29	C
ATOM	1180	CD	GLN	158A	-38.289	63.556	-52.084	1.00	14.62	C
ATOM	1181	OE1	GLN	158A	-38.784	62.949	-51.130	1.00	14.90	O
ATOM	1182	NE2	GLN	158A	-38.271	63.063	-53.319	1.00	15.65	N
ATOM	1183	C	GLN	158A	-36.681	67.140	-50.135	1.00	6.98	C
ATOM	1184	O	GLN	158A	-37.324	68.005	-50.718	1.00	6.78	O
ATOM	1185	N	GLU	159A	-36.771	66.932	-48.828	1.00	7.97	N
ATOM	1186	CA	GLU	159A	-37.649	67.733	-47.987	1.00	10.62	C
ATOM	1187	CB	GLU	159A	-37.589	67.222	-46.550	1.00	13.43	C
ATOM	1188	CG	GLU	159A	-38.298	68.084	-45.533	1.00	18.99	C
ATOM	1189	CD	GLU	159A	-38.485	67.362	-44.210	1.00	22.11	C
ATOM	1190	OE1	GLU	159A	-39.353	66.468	-44.143	1.00	23.07	O
ATOM	1191	OE2	GLU	159A	-37.760	67.678	-43.244	1.00	24.54	O
ATOM	1192	C	GLU	159A	-37.274	69.210	-48.050	1.00	8.96	C
ATOM	1193	O	GLU	159A	-38.150	70.066	-48.184	1.00	8.94	O
ATOM	1194	N	LEU	160A	-35.980	69.507	-47.962	1.00	8.71	N
ATOM	1195	CA	LEU	160A	-35.505	70.886	-48.021	1.00	7.21	C
ATOM	1196	CB	LEU	160A	-33.994	70.948	-47.764	1.00	7.92	C
ATOM	1197	CG	LEU	160A	-33.575	70.652	-46.322	1.00	9.01	C
ATOM	1198	CD1	LEU	160A	-32.057	70.639	-46.213	1.00	7.48	C
ATOM	1199	CD2	LEU	160A	-34.177	71.701	-45.388	1.00	10.34	C
ATOM	1200	C	LEU	160A	-35.835	71.532	-49.370	1.00	7.30	C
ATOM	1201	O	LEU	160A	-36.225	72.698	-49.419	1.00	6.22	O
ATOM	1202	N	LEU	161A	-35.677	70.778	-50.459	1.00	6.33	N
ATOM	1203	CA	LEU	161A	-35.990	71.304	-51.786	1.00	4.73	C
ATOM	1204	CB	LEU	161A	-35.638	70.288	-52.882	1.00	5.96	C
ATOM	1205	CG	LEU	161A	-36.132	70.640	-54.296	1.00	5.06	C
ATOM	1206	CD1	LEU	161A	-35.538	71.974	-54.753	1.00	3.20	C
ATOM	1207	CD2	LEU	161A	-35.744	69.526	-55.259	1.00	7.06	C
ATOM	1208	C	LEU	161A	-37.482	71.625	-51.841	1.00	3.16	C
ATOM	1209	O	LEU	161A	-37.880	72.687	-52.317	1.00	2.71	O
ATOM	1210	N	ASP	162A	-38.305	70.701	-51.353	1.00	2.50	N
ATOM	1211	CA	ASP	162A	-39.749	70.920	-51.335	1.00	5.20	C
ATOM	1212	CB	ASP	162A	-40.471	69.799	-50.584	1.00	4.63	C
ATOM	1213	CG	ASP	162A	-40.400	68.469	-51.293	1.00	9.21	C
ATOM	1214	OD1	ASP	162A	-40.235	68.439	-52.532	1.00	8.31	O
ATOM	1215	OD2	ASP	162A	-40.536	67.443	-50.597	1.00	9.33	O
ATOM	1216	C	ASP	162A	-40.072	72.232	-50.633	1.00	6.04	C
ATOM	1217	O	ASP	162A	-40.942	72.996	-51.074	1.00	5.18	O

FIGURE 9 - 20

ATOM	1218	N	GLU	163A	-39.374	72.477	-49.528	1.00	5.27	N
ATOM	1219	CA	GLU	163A	-39.582	73.679	-48.730	1.00	6.04	C
ATOM	1220	CB	GLU	163A	-38.774	73.583	-47.424	1.00	9.63	C
ATOM	1221	CG	GLU	163A	-39.178	72.401	-46.540	1.00	13.68	C
ATOM	1222	CD	GLU	163A	-38.327	72.256	-45.278	1.00	19.86	C
ATOM	1223	OE1	GLU	163A	-38.519	71.266	-44.541	1.00	21.59	O
ATOM	1224	OE2	GLU	163A	-37.471	73.125	-45.014	1.00	23.73	O
ATOM	1225	C	GLU	163A	-39.232	74.953	-49.494	1.00	6.07	C
ATOM	1226	O	GLU	163A	-39.989	75.925	-49.458	1.00	6.02	O
ATOM	1227	N	SER	164A	-38.100	74.951	-50.196	1.00	4.35	N
ATOM	1228	CA	SER	164A	-37.700	76.124	-50.971	1.00	4.11	C
ATOM	1229	CB	SER	164A	-36.280	75.956	-51.527	1.00	4.56	C
ATOM	1230	OG	SER	164A	-36.241	74.976	-52.550	1.00	6.81	O
ATOM	1231	C	SER	164A	-38.681	76.345	-52.126	1.00	4.17	C
ATOM	1232	O	SER	164A	-38.910	77.478	-52.552	1.00	4.06	O
ATOM	1233	N	ILE	165A	-39.250	75.257	-52.637	1.00	4.79	N
ATOM	1234	CA	ILE	165A	-40.219	75.348	-53.717	1.00	4.40	C
ATOM	1235	CB	ILE	165A	-40.579	73.933	-54.259	1.00	4.84	C
ATOM	1236	CG2	ILE	165A	-41.902	73.970	-55.026	1.00	3.37	C
ATOM	1237	CG1	ILE	165A	-39.436	73.434	-55.155	1.00	5.35	C
ATOM	1238	CD1	ILE	165A	-39.633	72.044	-55.732	1.00	5.21	C
ATOM	1239	C	ILE	165A	-41.455	76.049	-53.151	1.00	6.35	C
ATOM	1240	O	ILE	165A	-41.991	76.979	-53.757	1.00	5.68	O
ATOM	1241	N	ASP	166A	-41.886	75.625	-51.968	1.00	6.06	N
ATOM	1242	CA	ASP	166A	-43.043	76.252	-51.351	1.00	7.19	C
ATOM	1243	CB	ASP	166A	-43.424	75.541	-50.054	1.00	10.02	C
ATOM	1244	CG	ASP	166A	-44.725	76.061	-49.482	1.00	12.34	C
ATOM	1245	OD1	ASP	166A	-44.708	77.091	-48.775	1.00	16.98	O
ATOM	1246	OD2	ASP	166A	-45.770	75.452	-49.765	1.00	14.62	O
ATOM	1247	C	ASP	166A	-42.750	77.723	-51.080	1.00	6.93	C
ATOM	1248	O	ASP	166A	-43.634	78.572	-51.216	1.00	6.73	O
ATOM	1249	N	PHE	167A	-41.509	78.023	-50.697	1.00	6.69	N
ATOM	1250	CA	PHE	167A	-41.102	79.399	-50.442	1.00	8.73	C
ATOM	1251	CB	PHE	167A	-39.641	79.462	-49.976	1.00	11.67	C
ATOM	1252	CG	PHE	167A	-39.420	78.996	-48.565	1.00	15.63	C
ATOM	1253	CD1	PHE	167A	-38.127	78.863	-48.066	1.00	17.75	C
ATOM	1254	CD2	PHE	167A	-40.491	78.707	-47.726	1.00	18.91	C
ATOM	1255	CE1	PHE	167A	-37.904	78.453	-46.754	1.00	19.71	C
ATOM	1256	CE2	PHE	167A	-40.279	78.295	-46.410	1.00	20.80	C
ATOM	1257	CZ	PHE	167A	-38.984	78.169	-45.925	1.00	20.38	C
ATOM	1258	C	PHE	167A	-41.245	80.191	-51.735	1.00	6.86	C
ATOM	1259	O	PHE	167A	-41.707	81.329	-51.731	1.00	7.39	O
ATOM	1260	N	ALA	168A	-40.824	79.584	-52.841	1.00	6.07	N
ATOM	1261	CA	ALA	168A	-40.911	80.222	-54.153	1.00	4.85	C
ATOM	1262	CB	ALA	168A	-40.180	79.365	-55.202	1.00	6.74	C
ATOM	1263	C	ALA	168A	-42.369	80.447	-54.569	1.00	3.95	C
ATOM	1264	O	ALA	168A	-42.706	81.498	-55.115	1.00	2.50	O
ATOM	1265	N	LEU	169A	-43.235	79.466	-54.320	1.00	2.55	N
ATOM	1266	CA	LEU	169A	-44.640	79.620	-54.679	1.00	4.17	C
ATOM	1267	CB	LEU	169A	-45.428	78.330	-54.415	1.00	6.27	C
ATOM	1268	CG	LEU	169A	-45.379	77.190	-55.446	1.00	10.90	C
ATOM	1269	CD1	LEU	169A	-45.747	77.718	-56.828	1.00	11.75	C
ATOM	1270	CD2	LEU	169A	-44.006	76.574	-55.477	1.00	16.93	C
ATOM	1271	C	LEU	169A	-45.229	80.768	-53.870	1.00	2.04	C
ATOM	1272	O	LEU	169A	-46.162	81.436	-54.311	1.00	4.44	O
ATOM	1273	N	GLY	170A	-44.668	81.008	-52.688	1.00	3.89	N
ATOM	1274	CA	GLY	170A	-45.150	82.100	-51.854	1.00	3.81	C
ATOM	1275	C	GLY	170A	-44.989	83.469	-52.502	1.00	5.02	C
ATOM	1276	O	GLY	170A	-45.704	84.412	-52.158	1.00	6.69	O
ATOM	1277	N	VAL	171A	-44.043	83.602	-53.429	1.00	4.03	N
ATOM	1278	CA	VAL	171A	-43.848	84.885	-54.093	1.00	3.30	C

FIGURE 9 - 21

ATOM	1279	CB	VAL	171A	-42.356	85.329	-54.083	1.00	2.32	C
ATOM	1280	CG1	VAL	171A	-41.867	85.473	-52.654	1.00	1.00	C
ATOM	1281	CG2	VAL	171A	-41.493	84.326	-54.845	1.00	3.30	C
ATOM	1282	C	VAL	171A	-44.350	84.842	-55.534	1.00	2.80	C
ATOM	1283	O	VAL	171A	-43.896	85.604	-56.380	1.00	2.06	O
ATOM	1284	N	LEU	172A	-45.289	83.943	-55.812	1.00	1.74	N
ATOM	1285	CA	LEU	172A	-45.836	83.830	-57.155	1.00	2.02	C
ATOM	1286	CB	LEU	172A	-46.876	82.707	-57.225	1.00	2.04	C
ATOM	1287	CG	LEU	172A	-47.455	82.463	-58.618	1.00	3.25	C
ATOM	1288	CD1	LEU	172A	-46.328	82.116	-59.571	1.00	1.00	C
ATOM	1289	CD2	LEU	172A	-48.478	81.331	-58.577	1.00	3.39	C
ATOM	1290	C	LEU	172A	-46.470	85.148	-57.588	1.00	3.78	C
ATOM	1291	O	LEU	172A	-46.305	85.579	-58.732	1.00	2.36	O
ATOM	1292	N	PRO	173A	-47.223	85.801	-56.685	1.00	1.92	N
ATOM	1293	CD	PRO	173A	-47.691	85.350	-55.363	1.00	4.79	C
ATOM	1294	CA	PRO	173A	-47.848	87.074	-57.042	1.00	4.00	C
ATOM	1295	CB	PRO	173A	-48.461	87.532	-55.724	1.00	5.43	C
ATOM	1296	CG	PRO	173A	-48.885	86.259	-55.108	1.00	4.83	C
ATOM	1297	C	PRO	173A	-46.845	88.089	-57.587	1.00	3.75	C
ATOM	1298	O	PRO	173A	-47.121	88.790	-58.562	1.00	2.63	O
ATOM	1299	N	GLU	174A	-45.684	88.176	-56.946	1.00	3.22	N
ATOM	1300	CA	GLU	174A	-44.655	89.112	-57.380	1.00	4.83	C
ATOM	1301	CB	GLU	174A	-43.550	89.236	-56.322	1.00	6.97	C
ATOM	1302	CG	GLU	174A	-43.906	90.050	-55.076	1.00	7.95	C
ATOM	1303	CD	GLU	174A	-45.022	89.439	-54.255	1.00	8.81	C
ATOM	1304	OE1	GLU	174A	-45.183	88.199	-54.283	1.00	5.68	O
ATOM	1305	OE2	GLU	174A	-45.728	90.205	-53.563	1.00	7.94	O
ATOM	1306	C	GLU	174A	-44.032	88.663	-58.703	1.00	4.95	C
ATOM	1307	O	GLU	174A	-43.839	89.470	-59.609	1.00	6.20	O
ATOM	1308	N	MET	175A	-43.724	87.377	-58.823	1.00	5.12	N
ATOM	1309	CA	MET	175A	-43.109	86.897	-60.058	1.00	3.89	C
ATOM	1310	CB	MET	175A	-42.673	85.438	-59.905	1.00	2.78	C
ATOM	1311	CG	MET	175A	-41.725	85.212	-58.747	1.00	7.50	C
ATOM	1312	SD	MET	175A	-40.965	83.573	-58.770	1.00	7.54	S
ATOM	1313	CE	MET	175A	-42.374	82.552	-58.338	1.00	6.23	C
ATOM	1314	C	MET	175A	-44.028	87.052	-61.271	1.00	2.72	C
ATOM	1315	O	MET	175A	-43.587	87.482	-62.331	1.00	1.00	O
ATOM	1316	N	LEU	176A	-45.303	86.706	-61.129	1.00	3.41	N
ATOM	1317	CA	LEU	176A	-46.224	86.844	-62.251	1.00	3.01	C
ATOM	1318	CB	LEU	176A	-47.608	86.318	-61.871	1.00	4.67	C
ATOM	1319	CG	LEU	176A	-47.710	84.801	-61.716	1.00	5.70	C
ATOM	1320	CD1	LEU	176A	-49.133	84.451	-61.309	1.00	3.31	C
ATOM	1321	CD2	LEU	176A	-47.339	84.100	-63.026	1.00	3.06	C
ATOM	1322	C	LEU	176A	-46.332	88.298	-62.708	1.00	3.63	C
ATOM	1323	O	LEU	176A	-46.459	88.582	-63.901	1.00	1.32	O
ATOM	1324	N	ALA	177A	-46.269	89.217	-61.751	1.00	3.68	N
ATOM	1325	CA	ALA	177A	-46.361	90.639	-62.052	1.00	5.90	C
ATOM	1326	CB	ALA	177A	-46.693	91.418	-60.781	1.00	6.48	C
ATOM	1327	C	ALA	177A	-45.070	91.175	-62.656	1.00	5.56	C
ATOM	1328	O	ALA	177A	-45.073	92.215	-63.314	1.00	5.84	O
ATOM	1329	N	GLY	178A	-43.970	90.463	-62.428	1.00	6.30	N
ATOM	1330	CA	GLY	178A	-42.688	90.912	-62.936	1.00	5.58	C
ATOM	1331	C	GLY	178A	-42.007	91.808	-61.918	1.00	6.25	C
ATOM	1332	O	GLY	178A	-41.065	92.538	-62.243	1.00	6.12	O
ATOM	1333	N	ASP	179A	-42.485	91.751	-60.679	1.00	4.83	N
ATOM	1334	CA	ASP	179A	-41.927	92.557	-59.594	1.00	6.18	C
ATOM	1335	CB	ASP	179A	-43.031	92.890	-58.581	1.00	7.12	C
ATOM	1336	CG	ASP	179A	-42.653	94.024	-57.646	1.00	8.99	C
ATOM	1337	OD1	ASP	179A	-43.565	94.557	-56.977	1.00	8.23	O
ATOM	1338	OD2	ASP	179A	-41.455	94.374	-57.570	1.00	7.86	O
ATOM	1339	C	ASP	179A	-40.791	91.780	-58.923	1.00	6.13	C

FIGURE 9 - 22

```
ATOM   1340  O   ASP  179A   -40.885  91.357 -57.771  1.00  3.72 O
ATOM   1341  N   TRP  180A   -39.702  91.607 -59.659  1.00  7.35 N
ATOM   1342  CA  TRP  180A   -38.554  90.863 -59.164  1.00  7.76 C
ATOM   1343  CB  TRP  180A   -37.516  90.754 -60.278  1.00  7.58 C
ATOM   1344  CG  TRP  180A   -38.135  90.175 -61.506  1.00  8.22 C
ATOM   1345  CD2 TRP  180A   -38.608  88.832 -61.671  1.00  5.10 C
ATOM   1346  CE2 TRP  180A   -39.175  88.745 -62.957  1.00  3.86 C
ATOM   1347  CE3 TRP  180A   -38.605  87.692 -60.851  1.00  5.18 C
ATOM   1348  CD1 TRP  180A   -38.428  90.829 -62.666  1.00  6.12 C
ATOM   1349  NE1 TRP  180A   -39.053  89.977 -63.543  1.00  6.40 N
ATOM   1350  CZ2 TRP  180A   -39.736  87.564 -63.450  1.00  3.62 C
ATOM   1351  CZ3 TRP  180A   -39.166  86.513 -61.342  1.00  2.30 C
ATOM   1352  CH2 TRP  180A   -39.722  86.460 -62.631  1.00  1.00 C
ATOM   1353  C   TRP  180A   -37.944  91.451 -57.903  1.00  8.98 C
ATOM   1354  O   TRP  180A   -37.550  90.719 -56.996  1.00  9.71 O
ATOM   1355  N   THR  181A   -37.875  92.773 -57.843  1.00 10.45 N
ATOM   1356  CA  THR  181A   -37.323  93.448 -56.679  1.00 10.67 C
ATOM   1357  CB  THR  181A   -37.522  94.959 -56.797  1.00 13.56 C
ATOM   1358  OG1 THR  181A   -38.922  95.241 -56.895  1.00 18.11 O
ATOM   1359  CG2 THR  181A   -36.829  95.484 -58.041  1.00 13.17 C
ATOM   1360  C   THR  181A   -38.017  92.953 -55.410  1.00  9.05 C
ATOM   1361  O   THR  181A   -37.365  92.510 -54.465  1.00  7.65 O
ATOM   1362  N   ARG  182A   -39.346  93.022 -55.404  1.00  9.67 N
ATOM   1363  CA  ARG  182A   -40.140  92.586 -54.259  1.00  8.86 C
ATOM   1364  CB  ARG  182A   -41.604  92.956 -54.482  1.00 12.32 C
ATOM   1365  CG  ARG  182A   -42.518  92.649 -53.312  1.00 16.11 C
ATOM   1366  CD  ARG  182A   -42.265  93.565 -52.132  1.00 22.02 C
ATOM   1367  NE  ARG  182A   -43.349  93.464 -51.159  1.00 29.18 N
ATOM   1368  CZ  ARG  182A   -43.477  94.242 -50.090  1.00 31.14 C
ATOM   1369  NH1 ARG  182A   -44.502  94.071 -49.266  1.00 32.94 N
ATOM   1370  NH2 ARG  182A   -42.583  95.192 -49.847  1.00 32.49 N
ATOM   1371  C   ARG  182A   -40.024  91.076 -54.032  1.00  8.94 C
ATOM   1372  O   ARG  182A   -39.887  90.614 -52.898  1.00  7.80 O
ATOM   1373  N   ALA  183A   -40.094  90.312 -55.113  1.00  7.15 N
ATOM   1374  CA  ALA  183A   -39.992  88.862 -55.017  1.00  8.70 C
ATOM   1375  CB  ALA  183A   -40.081  88.236 -56.405  1.00  9.10 C
ATOM   1376  C   ALA  183A   -38.678  88.475 -54.350  1.00  8.80 C
ATOM   1377  O   ALA  183A   -38.648  87.643 -53.442  1.00  7.93 O
ATOM   1378  N   MET  184A   -37.594  89.099 -54.795  1.00 10.12 N
ATOM   1379  CA  MET  184A   -36.281  88.815 -54.247  1.00 10.88 C
ATOM   1380  CB  MET  184A   -35.207  89.540 -55.060  1.00 11.54 C
ATOM   1381  CG  MET  184A   -35.078  89.004 -56.478  1.00 14.15 C
ATOM   1382  SD  MET  184A   -33.816  89.825 -57.464  1.00 18.18 S
ATOM   1383  CE  MET  184A   -32.389  88.937 -56.939  1.00 17.44 C
ATOM   1384  C   MET  184A   -36.148  89.154 -52.765  1.00 11.97 C
ATOM   1385  O   MET  184A   -35.515  88.416 -52.020  1.00 12.18 O
ATOM   1386  N   GLN  185A   -36.742  90.256 -52.322  1.00 12.75 N
ATOM   1387  CA  GLN  185A   -36.628  90.601 -50.910  1.00 15.73 C
ATOM   1388  CB  GLN  185A   -37.011  92.065 -50.688  1.00 18.93 C
ATOM   1389  CG  GLN  185A   -36.159  93.022 -51.510  1.00 26.45 C
ATOM   1390  CD  GLN  185A   -34.688  92.617 -51.538  1.00 30.92 C
ATOM   1391  OE1 GLN  185A   -34.030  92.533 -50.498  1.00 33.71 O
ATOM   1392  NE2 GLN  185A   -34.168  92.362 -52.737  1.00 32.42 N
ATOM   1393  C   GLN  185A   -37.468  89.681 -50.028  1.00 14.71 C
ATOM   1394  O   GLN  185A   -37.107  89.398 -48.886  1.00 14.60 O
ATOM   1395  N   LYS  186A   -38.583  89.198 -50.561  1.00 13.04 N
ATOM   1396  CA  LYS  186A   -39.438  88.304 -49.796  1.00 13.76 C
ATOM   1397  CB  LYS  186A   -40.857  88.289 -50.379  1.00 13.45 C
ATOM   1398  CG  LYS  186A   -41.539  89.655 -50.367  1.00 16.00 C
ATOM   1399  CD  LYS  186A   -42.959  89.611 -50.908  1.00 17.17 C
ATOM   1400  CE  LYS  186A   -43.914  88.937 -49.934  1.00 19.26 C
```

FIGURE 9 - 23

```
ATOM   1401  NZ  LYS   186A     -45.338  89.130 -50.333  1.00 19.39 N
ATOM   1402  C   LYS   186A     -38.858  86.894 -49.808  1.00 14.50 C
ATOM   1403  O   LYS   186A     -38.763  86.244 -48.767  1.00 12.91 O
ATOM   1404  N   LEU   187A     -38.453  86.437 -50.990  1.00 14.43 N
ATOM   1405  CA  LEU   187A     -37.899  85.098 -51.147  1.00 16.13 C
ATOM   1406  CB  LEU   187A     -37.796  84.747 -52.631  1.00 13.71 C
ATOM   1407  CG  LEU   187A     -37.121  83.420 -52.985  1.00 12.09 C
ATOM   1408  CD1 LEU   187A     -37.845  82.270 -52.305  1.00  8.66 C
ATOM   1409  CD2 LEU   187A     -37.119  83.237 -54.496  1.00  7.97 C
ATOM   1410  C   LEU   187A     -36.539  84.908 -50.493  1.00 18.49 C
ATOM   1411  O   LEU   187A     -36.348  83.988 -49.698  1.00 20.01 O
ATOM   1412  N   HIS   188A     -35.593  85.774 -50.839  1.00 21.06 N
ATOM   1413  CA  HIS   188A     -34.239  85.687 -50.306  1.00 23.31 C
ATOM   1414  CB  HIS   188A     -33.366  86.777 -50.933  1.00 21.50 C
ATOM   1415  CG  HIS   188A     -33.161  86.614 -52.409  1.00 18.52 C
ATOM   1416  CD2 HIS   188A     -32.320  87.242 -53.265  1.00 17.87 C
ATOM   1417  ND1 HIS   188A     -33.894  85.730 -53.171  1.00 18.23 N
ATOM   1418  CE1 HIS   188A     -33.514  85.821 -54.434  1.00 15.96 C
ATOM   1419  NE2 HIS   188A     -32.562  86.732 -54.517  1.00 17.70 N
ATOM   1420  C   HIS   188A     -34.170  85.780 -48.783  1.00 26.12 C
ATOM   1421  O   HIS   188A     -33.149  85.443 -48.185  1.00 27.38 O
ATOM   1422  N   SER   189A     -35.253  86.230 -48.158  1.00 29.36 N
ATOM   1423  CA  SER   189A     -35.286  86.359 -46.703  1.00 32.73 C
ATOM   1424  CB  SER   189A     -36.080  87.604 -46.302  1.00 33.66 C
ATOM   1425  OG  SER   189A     -37.427  87.506 -46.732  1.00 35.88 O
ATOM   1426  C   SER   189A     -35.900  85.133 -46.034  1.00 33.46 C
ATOM   1427  O   SER   189A     -36.105  85.118 -44.822  1.00 32.82 O
ATOM   1428  N   GLN   190A     -36.189  84.107 -46.827  1.00 34.94 N
ATOM   1429  CA  GLN   190A     -36.782  82.877 -46.309  1.00 37.56 C
ATOM   1430  CB  GLN   190A     -37.712  82.259 -47.354  1.00 37.62 C
ATOM   1431  CG  GLN   190A     -38.893  83.133 -47.730  1.00 38.10 C
ATOM   1432  CD  GLN   190A     -39.755  83.480 -46.536  1.00 38.82 C
ATOM   1433  OE1 GLN   190A     -40.284  82.596 -45.858  1.00 37.83 O
ATOM   1434  NE2 GLN   190A     -39.900  84.774 -46.267  1.00 38.91 N
ATOM   1435  C   GLN   190A     -35.712  81.862 -45.924  1.00 38.68 C
ATOM   1436  O   GLN   190A     -34.851  81.517 -46.734  1.00 38.29 O
ATOM   1437  N   LYS   191A     -35.775  81.384 -44.685  1.00 40.02 N
ATOM   1438  CA  LYS   191A     -34.814  80.407 -44.184  1.00 41.24 C
ATOM   1439  CB  LYS   191A     -34.915  79.100 -44.977  1.00 38.13 C
ATOM   1440  CG  LYS   191A     -35.246  77.878 -44.132  1.00 33.08 C
ATOM   1441  CD  LYS   191A     -34.185  77.628 -43.069  1.00 29.39 C
ATOM   1442  CE  LYS   191A     -34.557  76.454 -42.177  1.00 26.28 C
ATOM   1443  NZ  LYS   191A     -33.506  76.170 -41.156  1.00 23.08 N
ATOM   1444  C   LYS   191A     -33.391  80.945 -44.272  1.00 45.11 C
ATOM   1445  O   LYS   191A     -32.432  80.244 -43.944  1.00 44.97 O
ATOM   1446  N   ALA   192A     -33.260  82.189 -44.723  1.00 48.96 N
ATOM   1447  CA  ALA   192A     -31.956  82.832 -44.855  1.00 52.68 C
ATOM   1448  CB  ALA   192A     -31.831  83.487 -46.227  1.00 52.16 C
ATOM   1449  C   ALA   192A     -31.807  83.877 -43.755  1.00 55.30 C
ATOM   1450  O   ALA   192A     -31.417  85.018 -44.008  1.00 55.26 O
ATOM   1451  N   GLN   193A     -32.125  83.470 -42.530  1.00 58.43 N
ATOM   1452  CA  GLN   193A     -32.048  84.351 -41.373  1.00 61.30 C
ATOM   1453  CB  GLN   193A     -33.461  84.657 -40.871  1.00 62.98 C
ATOM   1454  CG  GLN   193A     -34.333  85.386 -41.883  1.00 65.13 C
ATOM   1455  CD  GLN   193A     -35.797  85.391 -41.493  1.00 66.25 C
ATOM   1456  OE1 GLN   193A     -36.158  85.809 -40.394  1.00 67.23 O
ATOM   1457  NE2 GLN   193A     -36.651  84.923 -42.395  1.00 66.93 N
ATOM   1458  C   GLN   193A     -31.225  83.716 -40.254  1.00 62.15 C
ATOM   1459  O   GLN   193A     -30.758  82.572 -40.440  1.00 62.90 O
ATOM   1460  OXT GLN   193A     -31.055  84.371 -39.205  1.00 63.09 O
TER    1461       GLN   193A
```

FIGURE 9 - 24

```
ATOM   1462  CB  THR   1B    -69.713  51.046 -75.665  1.00 14.17 C
ATOM   1463  OG1 THR   1B    -70.110  51.389 -74.332  1.00 19.10 O
ATOM   1464  CG2 THR   1B    -70.949  50.895 -76.548  1.00 19.35 C
ATOM   1465  C   THR   1B    -67.582  49.957 -74.944  1.00  7.70 C
ATOM   1466  O   THR   1B    -67.459  49.742 -73.738  1.00  4.35 O
ATOM   1467  N   THR   1B    -69.650  48.661 -75.010  1.00  8.19 N
ATOM   1468  CA  THR   1B    -68.912  49.735 -75.649  1.00 11.67 C
ATOM   1469  N   ILE   2B    -66.585  50.375 -75.716  1.00  6.24 N
ATOM   1470  CA  ILE   2B    -65.252  50.629 -75.190  1.00  7.18 C
ATOM   1471  CB  ILE   2B    -64.224  50.764 -76.345  1.00  9.65 C
ATOM   1472  CG2 ILE   2B    -64.696  51.784 -77.340  1.00 10.95 C
ATOM   1473  CG1 ILE   2B    -62.858  51.181 -75.807  1.00 11.66 C
ATOM   1474  CD1 ILE   2B    -62.136  50.072 -75.121  1.00 15.86 C
ATOM   1475  C   ILE   2B    -65.238  51.906 -74.359  1.00  5.63 C
ATOM   1476  O   ILE   2B    -65.788  52.932 -74.770  1.00  3.07 O
ATOM   1477  N   LYS   3B    -64.612  51.830 -73.189  1.00  3.42 N
ATOM   1478  CA  LYS   3B    -64.503  52.974 -72.293  1.00  5.41 C
ATOM   1479  CB  LYS   3B    -65.098  52.637 -70.928  1.00  9.75 C
ATOM   1480  CG  LYS   3B    -66.617  52.545 -70.916  1.00 15.41 C
ATOM   1481  CD  LYS   3B    -67.123  51.907 -69.625  1.00 20.90 C
ATOM   1482  CE  LYS   3B    -66.687  52.692 -68.399  1.00 23.21 C
ATOM   1483  NZ  LYS   3B    -67.036  51.980 -67.138  1.00 27.59 N
ATOM   1484  C   LYS   3B    -63.044  53.391 -72.117  1.00  2.74 C
ATOM   1485  O   LYS   3B    -62.756  54.472 -71.604  1.00  1.00 O
ATOM   1486  N   LEU   4B    -62.126  52.529 -72.540  1.00  3.03 N
ATOM   1487  CA  LEU   4B    -60.707  52.835 -72.411  1.00  1.99 C
ATOM   1488  CB  LEU   4B    -60.163  52.280 -71.091  1.00  1.00 C
ATOM   1489  CG  LEU   4B    -58.657  52.462 -70.895  1.00  1.30 C
ATOM   1490  CD1 LEU   4B    -58.348  53.950 -70.825  1.00  1.00 C
ATOM   1491  CD2 LEU   4B    -58.189  51.741 -69.628  1.00  1.00 C
ATOM   1492  C   LEU   4B    -59.881  52.281 -73.565  1.00  1.97 C
ATOM   1493  O   LEU   4B    -59.922  51.082 -73.860  1.00  1.00 O
ATOM   1494  N   ILE   5B    -59.119  53.157 -74.208  1.00  1.00 N
ATOM   1495  CA  ILE   5B    -58.257  52.750 -75.310  1.00  1.00 C
ATOM   1496  CB  ILE   5B    -58.602  53.492 -76.628  1.00  1.00 C
ATOM   1497  CG2 ILE   5B    -57.670  53.002 -77.744  1.00  1.31 C
ATOM   1498  CG1 ILE   5B    -60.055  53.210 -77.030  1.00  2.66 C
ATOM   1499  CD1 ILE   5B    -60.538  54.031 -78.233  1.00  4.42 C
ATOM   1500  C   ILE   5B    -56.826  53.067 -74.889  1.00  1.80 C
ATOM   1501  O   ILE   5B    -56.486  54.215 -74.611  1.00  1.00 O
ATOM   1502  N   VAL   6B    -56.002  52.025 -74.833  1.00  1.00 N
ATOM   1503  CA  VAL   6B    -54.613  52.140 -74.403  1.00  1.00 C
ATOM   1504  CB  VAL   6B    -54.290  51.067 -73.329  1.00  2.83 C
ATOM   1505  CG1 VAL   6B    -52.911  51.319 -72.753  1.00  1.00 C
ATOM   1506  CG2 VAL   6B    -55.340  51.069 -72.232  1.00  1.00 C
ATOM   1507  C   VAL   6B    -53.619  51.954 -75.555  1.00  1.00 C
ATOM   1508  O   VAL   6B    -53.769  51.041 -76.373  1.00  2.16 O
ATOM   1509  N   GLY   7B    -52.602  52.813 -75.597  1.00  1.82 N
ATOM   1510  CA  GLY   7B    -51.578  52.729 -76.624  1.00  1.00 C
ATOM   1511  C   GLY   7B    -50.235  52.481 -75.959  1.00  1.56 C
ATOM   1512  O   GLY   7B    -49.695  53.360 -75.282  1.00  1.00 O
ATOM   1513  N   LEU   8B    -49.698  51.282 -76.156  1.00  1.66 N
ATOM   1514  CA  LEU   8B    -48.431  50.883 -75.558  1.00  1.00 C
ATOM   1515  CB  LEU   8B    -48.331  49.355 -75.533  1.00  1.37 C
ATOM   1516  CG  LEU   8B    -49.472  48.691 -74.751  1.00  1.68 C
ATOM   1517  CD1 LEU   8B    -49.234  47.198 -74.637  1.00  1.00 C
ATOM   1518  CD2 LEU   8B    -49.570  49.303 -73.363  1.00  1.05 C
ATOM   1519  C   LEU   8B    -47.193  51.469 -76.222  1.00  2.78 C
ATOM   1520  O   LEU   8B    -47.143  51.655 -77.442  1.00  2.03 O
ATOM   1521  N   ALA   9B    -46.185  51.748 -75.402  1.00  2.33 N
ATOM   1522  CA  ALA   9B    -44.937  52.316 -75.893  1.00  4.92 C
```

FIGURE 9 - 25

```
ATOM   1523  CB  ALA   9B    -45.224  53.591 -76.654  1.00  8.45 C
ATOM   1524  C   ALA   9B    -44.027  52.624 -74.717  1.00  5.94 C
ATOM   1525  O   ALA   9B    -44.456  52.564 -73.564  1.00  9.56 O
ATOM   1526  N   ASN   10B   -42.771  52.942 -75.017  1.00  5.35 N
ATOM   1527  CA  ASN   10B   -41.799  53.298 -73.995  1.00  3.93 C
ATOM   1528  CB  ASN   10B   -40.445  52.641 -74.276  1.00  2.09 C
ATOM   1529  CG  ASN   10B   -40.378  51.204 -73.786  1.00  3.05 C
ATOM   1530  OD1 ASN   10B   -40.486  50.940 -72.588  1.00  3.44 O
ATOM   1531  ND2 ASN   10B   -40.182  50.269 -74.709  1.00  3.80 N
ATOM   1532  C   ASN   10B   -41.642  54.816 -74.011  1.00  4.25 C
ATOM   1533  O   ASN   10B   -41.789  55.450 -75.055  1.00  3.63 O
ATOM   1534  N   PRO   11B   -41.359  55.417 -72.846  1.00  4.95 N
ATOM   1535  CD  PRO   11B   -41.380  54.810 -71.502  1.00  5.59 C
ATOM   1536  CA  PRO   11B   -41.187  56.867 -72.756  1.00  6.89 C
ATOM   1537  CB  PRO   11B   -41.531  57.154 -71.306  1.00  6.22 C
ATOM   1538  CG  PRO   11B   -40.959  55.965 -70.612  1.00  5.48 C
ATOM   1539  C   PRO   11B   -39.763  57.291 -73.097  1.00  7.70 C
ATOM   1540  O   PRO   11B   -38.855  56.462 -73.129  1.00  8.71 O
ATOM   1541  N   GLY   12B   -39.585  58.581 -73.360  1.00  9.11 N
ATOM   1542  CA  GLY   12B   -38.264  59.102 -73.665  1.00  8.22 C
ATOM   1543  C   GLY   12B   -37.940  59.282 -75.134  1.00  7.20 C
ATOM   1544  O   GLY   12B   -38.426  58.535 -75.985  1.00  7.31 O
ATOM   1545  N   PRO   13B   -37.093  60.270 -75.458  1.00  6.58 N
ATOM   1546  CD  PRO   13B   -36.538  61.260 -74.514  1.00  5.90 C
ATOM   1547  CA  PRO   13B   -36.677  60.578 -76.825  1.00  7.02 C
ATOM   1548  CB  PRO   13B   -35.600  61.639 -76.616  1.00  7.12 C
ATOM   1549  CG  PRO   13B   -36.129  62.392 -75.434  1.00  4.41 C
ATOM   1550  C   PRO   13B   -36.169  59.391 -77.640  1.00  7.04 C
ATOM   1551  O   PRO   13B   -36.577  59.212 -78.784  1.00  7.74 O
ATOM   1552  N   GLU   14B   -35.299  58.580 -77.044  1.00  6.94 N
ATOM   1553  CA  GLU   14B   -34.717  57.434 -77.746  1.00  8.06 C
ATOM   1554  CB  GLU   14B   -33.560  56.841 -76.934  1.00 11.67 C
ATOM   1555  CG  GLU   14B   -33.068  57.735 -75.812  1.00 18.20 C
ATOM   1556  CD  GLU   14B   -33.949  57.647 -74.581  1.00 19.54 C
ATOM   1557  OE1 GLU   14B   -33.744  56.724 -73.765  1.00 22.49 O
ATOM   1558  OE2 GLU   14B   -34.854  58.486 -74.437  1.00 18.63 O
ATOM   1559  C   GLU   14B   -35.691  56.316 -78.113  1.00  6.25 C
ATOM   1560  O   GLU   14B   -35.374  55.477 -78.953  1.00  6.83 O
ATOM   1561  N   TYR   15B   -36.867  56.295 -77.497  1.00  2.91 N
ATOM   1562  CA  TYR   15B   -37.851  55.245 -77.785  1.00  5.46 C
ATOM   1563  CB  TYR   15B   -38.434  54.709 -76.478  1.00  5.06 C
ATOM   1564  CG  TYR   15B   -37.435  53.995 -75.591  1.00  5.88 C
ATOM   1565  CD1 TYR   15B   -37.029  52.696 -75.873  1.00  6.66 C
ATOM   1566  CE1 TYR   15B   -36.111  52.036 -75.055  1.00  7.09 C
ATOM   1567  CD2 TYR   15B   -36.899  54.625 -74.468  1.00  6.52 C
ATOM   1568  CE2 TYR   15B   -35.979  53.975 -73.643  1.00  7.59 C
ATOM   1569  CZ  TYR   15B   -35.593  52.680 -73.943  1.00  7.86 C
ATOM   1570  OH  TYR   15B   -34.711  52.021 -73.121  1.00  9.40 O
ATOM   1571  C   TYR   15B   -38.998  55.745 -78.653  1.00  6.23 C
ATOM   1572  O   TYR   15B   -39.791  54.962 -79.173  1.00  6.23 O
ATOM   1573  N   ASP   16B   -39.074  57.060 -78.811  1.00  8.89 N
ATOM   1574  CA  ASP   16B   -40.155  57.685 -79.557  1.00 10.79 C
ATOM   1575  CB  ASP   16B   -39.854  59.174 -79.731  1.00 11.71 C
ATOM   1576  CG  ASP   16B   -41.016  59.922 -80.325  1.00  9.63 C
ATOM   1577  OD1 ASP   16B   -42.164  59.609 -79.949  1.00 13.19 O
ATOM   1578  OD2 ASP   16B   -40.789  60.822 -81.153  1.00 10.28 O
ATOM   1579  C   ASP   16B   -40.630  57.097 -80.896  1.00 13.12 C
ATOM   1580  O   ASP   16B   -41.791  56.696 -81.010  1.00 16.26 O
ATOM   1581  N   GLN   17B   -39.766  57.038 -81.903  1.00 14.18 N
ATOM   1582  CA  GLN   17B   -40.185  56.518 -83.209  1.00 13.74 C
ATOM   1583  CB  GLN   17B   -39.536  57.326 -84.338  1.00 17.09 C
```

FIGURE 9 - 26

```
ATOM   1584  CG  GLN   17B    -39.745  58.829 -84.278  1.00 21.70 C
ATOM   1585  CD  GLN   17B    -39.194  59.524 -85.516  1.00 25.61 C
ATOM   1586  OE1 GLN   17B    -39.758  59.412 -86.608  1.00 27.66 O
ATOM   1587  NE2 GLN   17B    -38.083  60.233 -85.354  1.00 27.00 N
ATOM   1588  C   GLN   17B    -39.871  55.041 -83.443  1.00  8.93 C
ATOM   1589  O   GLN   17B    -39.773  54.599 -84.584  1.00 10.03 O
ATOM   1590  N   THR   18B    -39.722  54.275 -82.373  1.00  6.09 N
ATOM   1591  CA  THR   18B    -39.394  52.862 -82.517  1.00  3.46 C
ATOM   1592  CB  THR   18B    -38.623  52.364 -81.287  1.00  1.10 C
ATOM   1593  OG1 THR   18B    -39.432  52.518 -80.117  1.00  1.00 O
ATOM   1594  CG2 THR   18B    -37.330  53.169 -81.123  1.00  2.22 C
ATOM   1595  C   THR   18B    -40.628  51.997 -82.763  1.00  1.90 C
ATOM   1596  O   THR   18B    -41.752  52.388 -82.461  1.00  1.00 O
ATOM   1597  N   ARG   19B    -40.416  50.817 -83.329  1.00  1.60 N
ATOM   1598  CA  ARG   19B    -41.532  49.931 -83.640  1.00  1.45 C
ATOM   1599  CB  ARG   19B    -41.022  48.628 -84.281  1.00  1.38 C
ATOM   1600  CG  ARG   19B    -40.241  48.795 -85.593  1.00  1.00 C
ATOM   1601  CD  ARG   19B    -39.635  47.455 -86.019  1.00  2.24 C
ATOM   1602  NE  ARG   19B    -39.111  47.460 -87.381  1.00  3.25 N
ATOM   1603  CZ  ARG   19B    -37.820  47.466 -87.703  1.00  3.40 C
ATOM   1604  NH1 ARG   19B    -36.882  47.471 -86.764  1.00  5.26 N
ATOM   1605  NH2 ARG   19B    -37.466  47.454 -88.979  1.00  3.20 N
ATOM   1606  C   ARG   19B    -42.393  49.581 -82.427  1.00  1.00 C
ATOM   1607  O   ARG   19B    -43.616  49.498 -82.536  1.00  1.56 O
ATOM   1608  N   HIS   20B    -41.751  49.381 -81.280  1.00  1.00 N
ATOM   1609  CA  HIS   20B    -42.443  48.995 -80.049  1.00  1.00 C
ATOM   1610  CB  HIS   20B    -41.398  48.715 -78.961  1.00  3.27 C
ATOM   1611  CG  HIS   20B    -41.756  47.591 -78.036  1.00  2.12 C
ATOM   1612  CD2 HIS   20B    -41.487  47.400 -76.722  1.00  4.27 C
ATOM   1613  ND1 HIS   20B    -42.428  46.463 -78.453  1.00  1.93 N
ATOM   1614  CE1 HIS   20B    -42.559  45.628 -77.438  1.00  2.29 C
ATOM   1615  NE2 HIS   20B    -41.996  46.172 -76.376  1.00  1.00 N
ATOM   1616  C   HIS   20B    -43.468  50.036 -79.573  1.00  1.00 C
ATOM   1617  O   HIS   20B    -44.311  49.736 -78.735  1.00  1.00 O
ATOM   1618  N   ASN   21B    -43.394  51.253 -80.112  1.00  1.26 N
ATOM   1619  CA  ASN   21B    -44.325  52.322 -79.741  1.00  1.00 C
ATOM   1620  CB  ASN   21B    -43.593  53.666 -79.686  1.00  3.79 C
ATOM   1621  CG  ASN   21B    -42.853  53.867 -78.383  1.00  1.84 C
ATOM   1622  OD1 ASN   21B    -42.378  52.909 -77.774  1.00  1.00 O
ATOM   1623  ND2 ASN   21B    -42.757  55.113 -77.944  1.00  5.23 N
ATOM   1624  C   ASN   21B    -45.493  52.412 -80.718  1.00  1.76 C
ATOM   1625  O   ASN   21B    -46.187  53.423 -80.780  1.00  2.54 O
ATOM   1626  N   ALA   22B    -45.704  51.341 -81.477  1.00  3.67 N
ATOM   1627  CA  ALA   22B    -46.778  51.290 -82.461  1.00  1.00 C
ATOM   1628  CB  ALA   22B    -46.857  49.889 -83.064  1.00  1.00 C
ATOM   1629  C   ALA   22B    -48.134  51.682 -81.874  1.00  1.00 C
ATOM   1630  O   ALA   22B    -48.857  52.504 -82.450  1.00  4.52 O
ATOM   1631  N   GLY   23B    -48.462  51.093 -80.728  1.00  1.52 N
ATOM   1632  CA  GLY   23B    -49.732  51.352 -80.070  1.00  1.96 C
ATOM   1633  C   GLY   23B    -49.922  52.789 -79.632  1.00  2.74 C
ATOM   1634  O   GLY   23B    -50.999  53.361 -79.811  1.00  2.44 O
ATOM   1635  N   ALA   24B    -48.877  53.371 -79.049  1.00  1.00 N
ATOM   1636  CA  ALA   24B    -48.934  54.752 -78.592  1.00  1.25 C
ATOM   1637  CB  ALA   24B    -47.639  55.117 -77.881  1.00  1.00 C
ATOM   1638  C   ALA   24B    -49.162  55.677 -79.781  1.00  1.62 C
ATOM   1639  O   ALA   24B    -49.857  56.693 -79.668  1.00  1.62 O
ATOM   1640  N   LEU   25B    -48.574  55.328 -80.922  1.00  2.04 N
ATOM   1641  CA  LEU   25B    -48.737  56.134 -82.131  1.00  3.30 C
ATOM   1642  CB  LEU   25B    -47.872  55.575 -83.272  1.00  3.24 C
ATOM   1643  CG  LEU   25B    -47.965  56.306 -84.619  1.00  3.57 C
ATOM   1644  CD1 LEU   25B    -47.715  57.795 -84.442  1.00  4.11 C
```

FIGURE 9 - 27

```
ATOM   1645  CD2 LEU   25B    -46.946  55.713 -85.582  1.00  3.61 C
ATOM   1646  C   LEU   25B    -50.206  56.114 -82.532  1.00  3.74 C
ATOM   1647  O   LEU   25B    -50.771  57.135 -82.932  1.00  2.20 O
ATOM   1648  N   PHE   26B    -50.823  54.943 -82.419  1.00  3.68 N
ATOM   1649  CA  PHE   26B    -52.230  54.802 -82.752  1.00  3.09 C
ATOM   1650  CB  PHE   26B    -52.714  53.369 -82.506  1.00  2.74 C
ATOM   1651  CG  PHE   26B    -54.204  53.214 -82.637  1.00  2.57 C
ATOM   1652  CD1 PHE   26B    -54.818  53.339 -83.878  1.00  2.58 C
ATOM   1653  CD2 PHE   26B    -54.999  53.003 -81.516  1.00  3.15 C
ATOM   1654  CE1 PHE   26B    -56.206  53.262 -84.004  1.00  1.71 C
ATOM   1655  CE2 PHE   26B    -56.392  52.925 -81.631  1.00  4.52 C
ATOM   1656  CZ  PHE   26B    -56.992  53.055 -82.874  1.00  1.00 C
ATOM   1657  C   PHE   26B    -53.074  55.749 -81.902  1.00  2.36 C
ATOM   1658  O   PHE   26B    -53.903  56.487 -82.428  1.00  1.89 O
ATOM   1659  N   VAL   27B    -52.869  55.712 -80.587  1.00  1.00 N
ATOM   1660  CA  VAL   27B    -53.630  56.565 -79.679  1.00  1.00 C
ATOM   1661  CB  VAL   27B    -53.346  56.212 -78.211  1.00  1.00 C
ATOM   1662  CG1 VAL   27B    -54.162  57.122 -77.280  1.00  1.00 C
ATOM   1663  CG2 VAL   27B    -53.704  54.756 -77.964  1.00  1.00 C
ATOM   1664  C   VAL   27B    -53.327  58.037 -79.913  1.00  1.00 C
ATOM   1665  O   VAL   27B    -54.202  58.887 -79.756  1.00  1.00 O
ATOM   1666  N   GLU   28B    -52.092  58.342 -80.299  1.00  1.51 N
ATOM   1667  CA  GLU   28B    -51.732  59.725 -80.571  1.00  1.56 C
ATOM   1668  CB  GLU   28B    -50.220  59.851 -80.812  1.00  1.61 C
ATOM   1669  CG  GLU   28B    -49.777  61.235 -81.254  1.00  5.63 C
ATOM   1670  CD  GLU   28B    -48.362  61.567 -80.797  1.00  4.97 C
ATOM   1671  OE1 GLU   28B    -47.465  60.715 -80.950  1.00  5.58 O
ATOM   1672  OE2 GLU   28B    -48.149  62.683 -80.292  1.00 12.07 O
ATOM   1673  C   GLU   28B    -52.508  60.231 -81.783  1.00  3.14 C
ATOM   1674  O   GLU   28B    -52.994  61.364 -81.785  1.00  4.58 O
ATOM   1675  N   ARG   29B    -52.631  59.388 -82.805  1.00  1.00 N
ATOM   1676  CA  ARG   29B    -53.356  59.752 -84.021  1.00  2.54 C
ATOM   1677  CB  ARG   29B    -53.181  58.683 -85.099  1.00  4.85 C
ATOM   1678  CG  ARG   29B    -51.798  58.592 -85.708  1.00  7.81 C
ATOM   1679  CD  ARG   29B    -51.457  59.822 -86.526  1.00 10.14 C
ATOM   1680  NE  ARG   29B    -50.211  59.632 -87.261  1.00  9.57 N
ATOM   1681  CZ  ARG   29B    -49.675  60.524 -88.086  1.00  9.57 C
ATOM   1682  NH1 ARG   29B    -50.269  61.692 -88.292  1.00 10.92 N
ATOM   1683  NH2 ARG   29B    -48.549  60.235 -88.723  1.00 11.35 N
ATOM   1684  C   ARG   29B    -54.841  59.888 -83.730  1.00  1.00 C
ATOM   1685  O   ARG   29B    -55.494  60.842 -84.165  1.00  1.83 O
ATOM   1686  N   LEU   30B    -55.367  58.908 -83.005  1.00  1.51 N
ATOM   1687  CA  LEU   30B    -56.776  58.880 -82.636  1.00  2.09 C
ATOM   1688  CB  LEU   30B    -57.050  57.655 -81.765  1.00  2.62 C
ATOM   1689  CG  LEU   30B    -58.469  57.088 -81.674  1.00  1.09 C
ATOM   1690  CD1 LEU   30B    -58.575  56.297 -80.380  1.00  6.99 C
ATOM   1691  CD2 LEU   30B    -59.528  58.175 -81.710  1.00  4.33 C
ATOM   1692  C   LEU   30B    -57.125  60.142 -81.852  1.00  3.33 C
ATOM   1693  O   LEU   30B    -58.121  60.820 -82.138  1.00  1.00 O
ATOM   1694  N   ALA   31B    -56.289  60.458 -80.867  1.00  2.74 N
ATOM   1695  CA  ALA   31B    -56.518  61.625 -80.028  1.00  2.12 C
ATOM   1696  CB  ALA   31B    -55.438  61.723 -78.964  1.00  4.15 C
ATOM   1697  C   ALA   31B    -56.562  62.902 -80.857  1.00  2.84 C
ATOM   1698  O   ALA   31B    -57.458  63.721 -80.692  1.00  1.00 O
ATOM   1699  N   HIS   32B    -55.604  63.072 -81.761  1.00  2.81 N
ATOM   1700  CA  HIS   32B    -55.599  64.275 -82.578  1.00  3.23 C
ATOM   1701  CB  HIS   32B    -54.322  64.361 -83.420  1.00  4.61 C
ATOM   1702  CG  HIS   32B    -54.148  65.686 -84.091  1.00  7.18 C
ATOM   1703  CD2 HIS   32B    -53.778  66.889 -83.595  1.00  9.15 C
ATOM   1704  ND1 HIS   32B    -54.455  65.897 -85.417  1.00  8.21 N
ATOM   1705  CE1 HIS   32B    -54.284  67.173 -85.709  1.00  9.96 C
```

FIGURE 9 - 28

```
ATOM   1706  NE2 HIS   32B    -53.874  67.798 -84.620  1.00 10.09 N
ATOM   1707  C   HIS   32B    -56.832  64.323 -83.483  1.00  3.91 C
ATOM   1708  O   HIS   32B    -57.463  65.370 -83.630  1.00  4.32 O
ATOM   1709  N   ALA   33B    -57.180  63.183 -84.068  1.00  5.05 N
ATOM   1710  CA  ALA   33B    -58.333  63.108 -84.950  1.00  4.34 C
ATOM   1711  CB  ALA   33B    -58.388  61.738 -85.607  1.00  2.99 C
ATOM   1712  C   ALA   33B    -59.648  63.404 -84.224  1.00  6.03 C
ATOM   1713  O   ALA   33B    -60.606  63.879 -84.839  1.00  4.20 O
ATOM   1714  N   GLN   34B    -59.696  63.149 -82.918  1.00  5.65 N
ATOM   1715  CA  GLN   34B    -60.918  63.393 -82.153  1.00  7.73 C
ATOM   1716  CB  GLN   34B    -61.200  62.225 -81.201  1.00  6.98 C
ATOM   1717  CG  GLN   34B    -61.532  60.906 -81.884  1.00  9.48 C
ATOM   1718  CD  GLN   34B    -62.715  61.013 -82.833  1.00 11.18 C
ATOM   1719  OE1 GLN   34B    -63.739  61.610 -82.504  1.00 13.98 O
ATOM   1720  NE2 GLN   34B    -62.576  60.433 -84.016  1.00 13.79 N
ATOM   1721  C   GLN   34B    -60.888  64.697 -81.353  1.00  6.05 C
ATOM   1722  O   GLN   34B    -61.829  64.989 -80.608  1.00  4.70 O
ATOM   1723  N   GLY   35B    -59.810  65.461 -81.507  1.00  4.94 N
ATOM   1724  CA  GLY   35B    -59.669  66.727 -80.802  1.00  6.33 C
ATOM   1725  C   GLY   35B    -59.446  66.575 -79.306  1.00  7.40 C
ATOM   1726  O   GLY   35B    -59.812  67.455 -78.523  1.00  6.28 O
ATOM   1727  N   VAL   36B    -58.823  65.470 -78.907  1.00  4.97 N
ATOM   1728  CA  VAL   36B    -58.568  65.198 -77.497  1.00  4.13 C
ATOM   1729  CB  VAL   36B    -58.809  63.703 -77.174  1.00  3.38 C
ATOM   1730  CG1 VAL   36B    -58.464  63.414 -75.717  1.00  2.16 C
ATOM   1731  CG2 VAL   36B    -60.251  63.335 -77.477  1.00  3.47 C
ATOM   1732  C   VAL   36B    -57.156  65.543 -77.029  1.00  4.02 C
ATOM   1733  O   VAL   36B    -56.166  65.129 -77.639  1.00  4.44 O
ATOM   1734  N   SER   37B    -57.073  66.288 -75.933  1.00  4.82 N
ATOM   1735  CA  SER   37B    -55.785  66.655 -75.354  1.00  6.56 C
ATOM   1736  CB  SER   37B    -55.896  67.968 -74.568  1.00  8.88 C
ATOM   1737  OG  SER   37B    -56.163  69.069 -75.410  1.00 10.31 O
ATOM   1738  C   SER   37B    -55.347  65.554 -74.390  1.00  5.80 C
ATOM   1739  O   SER   37B    -56.124  65.132 -73.539  1.00  6.42 O
ATOM   1740  N   LEU   38B    -54.104  65.099 -74.509  1.00  5.09 N
ATOM   1741  CA  LEU   38B    -53.594  64.070 -73.608  1.00  5.24 C
ATOM   1742  CB  LEU   38B    -52.635  63.126 -74.346  1.00  7.30 C
ATOM   1743  CG  LEU   38B    -53.228  62.242 -75.447  1.00  8.28 C
ATOM   1744  CD1 LEU   38B    -54.372  61.410 -74.871  1.00  7.41 C
ATOM   1745  CD2 LEU   38B    -53.722  63.109 -76.595  1.00 12.19 C
ATOM   1746  C   LEU   38B    -52.863  64.757 -72.452  1.00  4.62 C
ATOM   1747  O   LEU   38B    -51.795  65.339 -72.640  1.00  5.24 O
ATOM   1748  N   VAL   39B    -53.444  64.688 -71.260  1.00  4.32 N
ATOM   1749  CA  VAL   39B    -52.850  65.319 -70.085  1.00  4.16 C
ATOM   1750  CB  VAL   39B    -53.947  65.818 -69.105  1.00  3.31 C
ATOM   1751  CG1 VAL   39B    -54.900  64.693 -68.767  1.00  8.08 C
ATOM   1752  CG2 VAL   39B    -53.316  66.343 -67.827  1.00  1.03 C
ATOM   1753  C   VAL   39B    -51.913  64.389 -69.323  1.00  2.61 C
ATOM   1754  O   VAL   39B    -52.225  63.213 -69.102  1.00  2.82 O
ATOM   1755  N   ALA   40B    -50.749  64.913 -68.952  1.00  2.98 N
ATOM   1756  CA  ALA   40B    -49.798  64.139 -68.172  1.00  1.00 C
ATOM   1757  CB  ALA   40B    -48.428  64.836 -68.142  1.00  1.00 C
ATOM   1758  C   ALA   40B    -50.403  64.095 -66.774  1.00  3.09 C
ATOM   1759  O   ALA   40B    -50.524  65.128 -66.109  1.00  4.80 O
ATOM   1760  N   ASP   41B    -50.812  62.908 -66.343  1.00  2.70 N
ATOM   1761  CA  ASP   41B    -51.398  62.746 -65.024  1.00  2.73 C
ATOM   1762  CB  ASP   41B    -52.880  62.367 -65.128  1.00  2.19 C
ATOM   1763  CG  ASP   41B    -53.615  62.547 -63.812  1.00  3.33 C
ATOM   1764  OD1 ASP   41B    -52.961  62.442 -62.758  1.00  4.07 O
ATOM   1765  OD2 ASP   41B    -54.841  62.785 -63.826  1.00  2.44 O
ATOM   1766  C   ASP   41B    -50.639  61.653 -64.281  1.00  2.67 C
```

FIGURE 9 - 29

```
ATOM   1767  O   ASP  41B    -50.751  60.470 -64.597  1.00  1.11 O
ATOM   1768  N   ARG  42B    -49.861  62.054 -63.288  1.00  5.24 N
ATOM   1769  CA  ARG  42B    -49.071  61.098 -62.528  1.00  5.50 C
ATOM   1770  CB  ARG  42B    -48.143  61.845 -61.580  1.00  5.82 C
ATOM   1771  CG  ARG  42B    -47.478  63.021 -62.262  1.00 11.39 C
ATOM   1772  CD  ARG  42B    -46.223  63.437 -61.546  1.00 14.28 C
ATOM   1773  NE  ARG  42B    -45.410  62.285 -61.213  1.00 14.25 N
ATOM   1774  CZ  ARG  42B    -44.098  62.323 -61.038  1.00 14.47 C
ATOM   1775  NH1 ARG  42B    -43.440  63.468 -61.167  1.00 11.91 N
ATOM   1776  NH2 ARG  42B    -43.449  61.211 -60.740  1.00 10.85 N
ATOM   1777  C   ARG  42B    -49.925  60.102 -61.758  1.00  3.65 C
ATOM   1778  O   ARG  42B    -49.463  59.007 -61.429  1.00  3.77 O
ATOM   1779  N   LYS  43B    -51.170  60.472 -61.474  1.00  3.86 N
ATOM   1780  CA  LYS  43B    -52.057  59.577 -60.748  1.00  4.12 C
ATOM   1781  CB  LYS  43B    -53.373  60.276 -60.407  1.00  4.21 C
ATOM   1782  CG  LYS  43B    -53.214  61.463 -59.473  1.00  5.72 C
ATOM   1783  CD  LYS  43B    -54.556  61.890 -58.908  1.00 10.20 C
ATOM   1784  CE  LYS  43B    -55.466  62.420 -59.992  1.00 10.59 C
ATOM   1785  NZ  LYS  43B    -54.933  63.701 -60.527  1.00 16.01 N
ATOM   1786  C   LYS  43B    -52.338  58.324 -61.564  1.00  3.86 C
ATOM   1787  O   LYS  43B    -52.743  57.300 -61.017  1.00  4.67 O
ATOM   1788  N   TYR  44B    -52.112  58.400 -62.873  1.00  4.13 N
ATOM   1789  CA  TYR  44B    -52.361  57.251 -63.744  1.00  3.46 C
ATOM   1790  CB  TYR  44B    -53.511  57.569 -64.694  1.00  2.46 C
ATOM   1791  CG  TYR  44B    -54.730  58.049 -63.946  1.00  4.57 C
ATOM   1792  CD1 TYR  44B    -55.114  59.384 -63.991  1.00  6.35 C
ATOM   1793  CE1 TYR  44B    -56.202  59.848 -63.253  1.00  7.49 C
ATOM   1794  CD2 TYR  44B    -55.466  57.174 -63.145  1.00  5.10 C
ATOM   1795  CE2 TYR  44B    -56.554  57.623 -62.404  1.00  5.75 C
ATOM   1796  CZ  TYR  44B    -56.915  58.962 -62.464  1.00  9.05 C
ATOM   1797  OH  TYR  44B    -57.997  59.408 -61.743  1.00  9.39 O
ATOM   1798  C   TYR  44B    -51.128  56.822 -64.525  1.00  1.98 C
ATOM   1799  O   TYR  44B    -51.222  56.107 -65.523  1.00  1.00 O
ATOM   1800  N   PHE  45B    -49.969  57.276 -64.060  1.00  2.20 N
ATOM   1801  CA  PHE  45B    -48.697  56.911 -64.674  1.00  1.00 C
ATOM   1802  CB  PHE  45B    -48.322  55.486 -64.240  1.00  2.60 C
ATOM   1803  CG  PHE  45B    -48.355  55.283 -62.749  1.00  3.03 C
ATOM   1804  CD1 PHE  45B    -47.352  55.808 -61.937  1.00  5.79 C
ATOM   1805  CD2 PHE  45B    -49.415  54.617 -62.149  1.00  4.76 C
ATOM   1806  CE1 PHE  45B    -47.412  55.676 -60.550  1.00  6.82 C
ATOM   1807  CE2 PHE  45B    -49.483  54.480 -60.760  1.00  6.40 C
ATOM   1808  CZ  PHE  45B    -48.481  55.012 -59.961  1.00  5.39 C
ATOM   1809  C   PHE  45B    -48.742  57.003 -66.197  1.00  2.63 C
ATOM   1810  O   PHE  45B    -48.285  56.097 -66.898  1.00  2.11 O
ATOM   1811  N   GLY  46B    -49.290  58.099 -66.712  1.00  1.00 N
ATOM   1812  CA  GLY  46B    -49.343  58.245 -68.152  1.00  3.06 C
ATOM   1813  C   GLY  46B    -50.063  59.480 -68.642  1.00  2.11 C
ATOM   1814  O   GLY  46B    -50.411  60.360 -67.859  1.00  1.94 O
ATOM   1815  N   LEU  47B    -50.273  59.528 -69.953  1.00  2.30 N
ATOM   1816  CA  LEU  47B    -50.962  60.625 -70.615  1.00  1.51 C
ATOM   1817  CB  LEU  47B    -50.349  60.862 -71.987  1.00  1.00 C
ATOM   1818  CG  LEU  47B    -48.825  61.008 -72.018  1.00  2.31 C
ATOM   1819  CD1 LEU  47B    -48.375  61.271 -73.452  1.00  2.89 C
ATOM   1820  CD2 LEU  47B    -48.387  62.134 -71.100  1.00  1.00 C
ATOM   1821  C   LEU  47B    -52.406  60.184 -70.779  1.00  1.30 C
ATOM   1822  O   LEU  47B    -52.664  59.146 -71.377  1.00  2.38 O
ATOM   1823  N   VAL  48B    -53.345  60.970 -70.263  1.00  2.03 N
ATOM   1824  CA  VAL  48B    -54.750  60.593 -70.342  1.00  2.91 C
ATOM   1825  CB  VAL  48B    -55.308  60.272 -68.937  1.00  5.58 C
ATOM   1826  CG1 VAL  48B    -54.454  59.193 -68.272  1.00  7.03 C
ATOM   1827  CG2 VAL  48B    -55.342  61.525 -68.085  1.00 10.71 C
```

FIGURE 9 -30

```
ATOM   1828  C   VAL  48B   -55.649  61.632 -70.994  1.00  2.78 C
ATOM   1829  O   VAL  48B   -55.549  62.821 -70.714  1.00  1.09 O
ATOM   1830  N   GLY  49B   -56.528  61.160 -71.870  1.00  3.32 N
ATOM   1831  CA  GLY  49B   -57.453  62.041 -72.556  1.00  4.17 C
ATOM   1832  C   GLY  49B   -58.874  61.525 -72.427  1.00  4.82 C
ATOM   1833  O   GLY  49B   -59.097  60.373 -72.064  1.00  1.00 O
ATOM   1834  N   LYS  50B   -59.849  62.363 -72.734  1.00  5.07 N
ATOM   1835  CA  LYS  50B   -61.230  61.922 -72.614  1.00  7.20 C
ATOM   1836  CB  LYS  50B   -61.723  62.184 -71.185  1.00 10.55 C
ATOM   1837  CG  LYS  50B   -63.106  61.644 -70.876  1.00 16.38 C
ATOM   1838  CD  LYS  50B   -63.554  62.041 -69.474  1.00 20.00 C
ATOM   1839  CE  LYS  50B   -63.640  63.555 -69.323  1.00 23.07 C
ATOM   1840  NZ  LYS  50B   -64.114  63.960 -67.969  1.00 25.94 N
ATOM   1841  C   LYS  50B   -62.134  62.631 -73.605  1.00  5.78 C
ATOM   1842  O   LYS  50B   -61.916  63.797 -73.919  1.00  3.76 O
ATOM   1843  N   PHE  51B   -63.127  61.906 -74.114  1.00  6.08 N
ATOM   1844  CA  PHE  51B   -64.109  62.470 -75.030  1.00  7.04 C
ATOM   1845  CB  PHE  51B   -63.656  62.401 -76.501  1.00  8.44 C
ATOM   1846  CG  PHE  51B   -63.477  61.012 -77.042  1.00  9.37 C
ATOM   1847  CD1 PHE  51B   -62.389  60.234 -76.664  1.00 11.86 C
ATOM   1848  CD2 PHE  51B   -64.375  60.500 -77.981  1.00 11.25 C
ATOM   1849  CE1 PHE  51B   -62.190  58.965 -77.213  1.00 10.65 C
ATOM   1850  CE2 PHE  51B   -64.188  59.233 -78.537  1.00 11.95 C
ATOM   1851  CZ  PHE  51B   -63.090  58.463 -78.153  1.00 12.36 C
ATOM   1852  C   PHE  51B   -65.419  61.727 -74.832  1.00  8.35 C
ATOM   1853  O   PHE  51B   -65.464  60.707 -74.143  1.00  5.80 O
ATOM   1854  N   SER  52B   -66.488  62.244 -75.420  1.00  7.02 N
ATOM   1855  CA  SER  52B   -67.782  61.611 -75.264  1.00  8.60 C
ATOM   1856  CB  SER  52B   -68.837  62.655 -74.892  1.00 11.89 C
ATOM   1857  OG  SER  52B   -68.501  63.313 -73.686  1.00 16.40 O
ATOM   1858  C   SER  52B   -68.233  60.889 -76.512  1.00  7.29 C
ATOM   1859  O   SER  52B   -67.967  61.329 -77.624  1.00  7.38 O
ATOM   1860  N   HIS  53B   -68.916  59.767 -76.318  1.00  6.91 N
ATOM   1861  CA  HIS  53B   -69.451  59.007 -77.427  1.00  4.00 C
ATOM   1862  CB  HIS  53B   -68.468  57.969 -77.939  1.00  6.50 C
ATOM   1863  CG  HIS  53B   -68.874  57.385 -79.253  1.00  4.95 C
ATOM   1864  CD2 HIS  53B   -68.974  57.950 -80.480  1.00  2.78 C
ATOM   1865  ND1 HIS  53B   -69.283  56.078 -79.398  1.00  4.84 N
ATOM   1866  CE1 HIS  53B   -69.615  55.862 -80.658  1.00  4.28 C
ATOM   1867  NE2 HIS  53B   -69.437  56.981 -81.335  1.00  5.11 N
ATOM   1868  C   HIS  53B   -70.720  58.307 -76.989  1.00  5.23 C
ATOM   1869  O   HIS  53B   -70.710  57.521 -76.044  1.00  2.21 O
ATOM   1870  N   GLN  54B   -71.809  58.596 -77.691  1.00  5.12 N
ATOM   1871  CA  GLN  54B   -73.106  58.013 -77.363  1.00  5.26 C
ATOM   1872  CB  GLN  54B   -73.153  56.531 -77.770  1.00  4.22 C
ATOM   1873  CG  GLN  54B   -73.133  56.350 -79.285  1.00  4.12 C
ATOM   1874  CD  GLN  54B   -73.215  54.902 -79.741  1.00  5.05 C
ATOM   1875  OE1 GLN  54B   -73.112  54.620 -80.933  1.00  6.32 O
ATOM   1876  NE2 GLN  54B   -73.409  53.982 -78.801  1.00  4.68 N
ATOM   1877  C   GLN  54B   -73.412  58.184 -75.877  1.00  5.51 C
ATOM   1878  O   GLN  54B   -73.874  57.256 -75.203  1.00  8.09 O
ATOM   1879  N   GLY  55B   -73.144  59.390 -75.383  1.00  6.38 N
ATOM   1880  CA  GLY  55B   -73.416  59.731 -73.998  1.00  9.07 C
ATOM   1881  C   GLY  55B   -72.532  59.131 -72.925  1.00  8.91 C
ATOM   1882  O   GLY  55B   -72.811  59.308 -71.741  1.00  8.26 O
ATOM   1883  N   ALA  56B   -71.479  58.421 -73.320  1.00 10.25 N
ATOM   1884  CA  ALA  56B   -70.574  57.811 -72.352  1.00  9.81 C
ATOM   1885  CB  ALA  56B   -70.558  56.298 -72.525  1.00  9.66 C
ATOM   1886  C   ALA  56B   -69.163  58.373 -72.509  1.00  9.36 C
ATOM   1887  O   ALA  56B   -68.784  58.834 -73.586  1.00  9.03 O
ATOM   1888  N   ASP  57B   -68.397  58.344 -71.423  1.00 10.54 N
```

FIGURE 9 - 31

```
ATOM   1889  CA  ASP   57B    -67.023  58.838 -71.424  1.00  8.91 C
ATOM   1890  CB  ASP   57B    -66.601  59.241 -70.007  1.00 11.75 C
ATOM   1891  CG  ASP   57B    -67.172  60.583 -69.576  1.00 14.68 C
ATOM   1892  OD1 ASP   57B    -67.158  60.863 -68.361  1.00 16.93 O
ATOM   1893  OD2 ASP   57B    -67.612  61.364 -70.443  1.00 17.83 O
ATOM   1894  C   ASP   57B    -66.063  57.766 -71.929  1.00  8.22 C
ATOM   1895  O   ASP   57B    -66.042  56.656 -71.399  1.00  8.19 O
ATOM   1896  N   VAL   58B    -65.278  58.100 -72.952  1.00  5.00 N
ATOM   1897  CA  VAL   58B    -64.291  57.175 -73.504  1.00  4.89 C
ATOM   1898  CB  VAL   58B    -64.487  56.974 -75.029  1.00  2.61 C
ATOM   1899  CG1 VAL   58B    -63.493  55.945 -75.558  1.00  2.22 C
ATOM   1900  CG2 VAL   58B    -65.912  56.526 -75.312  1.00  1.51 C
ATOM   1901  C   VAL   58B    -62.906  57.775 -73.249  1.00  4.92 C
ATOM   1902  O   VAL   58B    -62.611  58.872 -73.711  1.00  5.02 O
ATOM   1903  N   ARG   59B    -62.059  57.058 -72.517  1.00  3.99 N
ATOM   1904  CA  ARG   59B    -60.727  57.570 -72.207  1.00  4.16 C
ATOM   1905  CB  ARG   59B    -60.440  57.395 -70.708  1.00  3.88 C
ATOM   1906  CG  ARG   59B    -61.364  58.252 -69.840  1.00  5.88 C
ATOM   1907  CD  ARG   59B    -61.009  58.248 -68.356  1.00  5.75 C
ATOM   1908  NE  ARG   59B    -61.755  59.286 -67.636  1.00  9.77 N
ATOM   1909  CZ  ARG   59B    -63.070  59.262 -67.426  1.00 11.45 C
ATOM   1910  NH1 ARG   59B    -63.663  60.254 -66.768  1.00 10.65 N
ATOM   1911  NH2 ARG   59B    -63.798  58.241 -67.864  1.00  9.53 N
ATOM   1912  C   ARG   59B    -59.593  56.974 -73.033  1.00  2.95 C
ATOM   1913  O   ARG   59B    -59.634  55.805 -73.434  1.00  1.00 O
ATOM   1914  N   LEU   60B    -58.597  57.816 -73.297  1.00  2.31 N
ATOM   1915  CA  LEU   60B    -57.418  57.441 -74.062  1.00  1.50 C
ATOM   1916  CB  LEU   60B    -57.202  58.431 -75.209  1.00  1.00 C
ATOM   1917  CG  LEU   60B    -58.380  58.575 -76.179  1.00  2.84 C
ATOM   1918  CD1 LEU   60B    -58.042  59.561 -77.278  1.00  4.61 C
ATOM   1919  CD2 LEU   60B    -58.704  57.207 -76.769  1.00  1.72 C
ATOM   1920  C   LEU   60B    -56.220  57.454 -73.116  1.00  2.09 C
ATOM   1921  O   LEU   60B    -56.018  58.407 -72.360  1.00  1.82 O
ATOM   1922  N   LEU   61B    -55.421  56.397 -73.160  1.00  1.00 N
ATOM   1923  CA  LEU   61B    -54.276  56.301 -72.271  1.00  1.00 C
ATOM   1924  CB  LEU   61B    -54.595  55.336 -71.121  1.00  2.58 C
ATOM   1925  CG  LEU   61B    -53.412  54.866 -70.262  1.00  2.10 C
ATOM   1926  CD1 LEU   61B    -52.891  56.015 -69.410  1.00  3.60 C
ATOM   1927  CD2 LEU   61B    -53.857  53.697 -69.387  1.00  5.37 C
ATOM   1928  C   LEU   61B    -52.990  55.853 -72.928  1.00  1.03 C
ATOM   1929  O   LEU   61B    -52.961  54.875 -73.677  1.00  2.58 O
ATOM   1930  N   ILE   62B    -51.921  56.591 -72.648  1.00  1.00 N
ATOM   1931  CA  ILE   62B    -50.605  56.219 -73.134  1.00  1.60 C
ATOM   1932  CB  ILE   62B    -49.993  57.253 -74.094  1.00  1.65 C
ATOM   1933  CG2 ILE   62B    -48.575  56.830 -74.463  1.00  2.13 C
ATOM   1934  CG1 ILE   62B    -50.833  57.344 -75.368  1.00  4.34 C
ATOM   1935  CD1 ILE   62B    -50.240  58.272 -76.411  1.00  4.19 C
ATOM   1936  C   ILE   62B    -49.756  56.149 -71.882  1.00  1.00 C
ATOM   1937  O   ILE   62B    -49.363  57.180 -71.331  1.00  2.27 O
ATOM   1938  N   PRO   63B    -49.500  54.930 -71.383  1.00  1.00 N
ATOM   1939  CD  PRO   63B    -49.978  53.624 -71.865  1.00  2.24 C
ATOM   1940  CA  PRO   63B    -48.682  54.775 -70.176  1.00  1.00 C
ATOM   1941  CB  PRO   63B    -48.591  53.259 -70.015  1.00  2.21 C
ATOM   1942  CG  PRO   63B    -49.872  52.771 -70.617  1.00  1.00 C
ATOM   1943  C   PRO   63B    -47.306  55.394 -70.402  1.00  1.74 C
ATOM   1944  O   PRO   63B    -46.777  55.329 -71.511  1.00  1.00 O
ATOM   1945  N   THR   64B    -46.730  55.983 -69.355  1.00  1.00 N
ATOM   1946  CA  THR   64B    -45.403  56.583 -69.450  1.00  2.14 C
ATOM   1947  CB  THR   64B    -45.422  58.085 -69.102  1.00  3.01 C
ATOM   1948  OG1 THR   64B    -45.980  58.271 -67.794  1.00  4.59 O
ATOM   1949  CG2 THR   64B    -46.246  58.849 -70.132  1.00  2.35 C
```

FIGURE 9 - 32

ATOM	1950	C	THR	64B	-44.411	55.861	-68.543	1.00	1.68	C
ATOM	1951	O	THR	64B	-43.310	56.348	-68.285	1.00	2.15	O
ATOM	1952	N	THR	65B	-44.830	54.702	-68.045	1.00	2.52	N
ATOM	1953	CA	THR	65B	-43.968	53.869	-67.230	1.00	3.10	C
ATOM	1954	CB	THR	65B	-44.771	52.818	-66.447	1.00	4.88	C
ATOM	1955	OG1	THR	65B	-45.573	52.064	-67.367	1.00	6.05	O
ATOM	1956	CG2	THR	65B	-45.674	53.475	-65.422	1.00	3.38	C
ATOM	1957	C	THR	65B	-43.227	53.139	-68.335	1.00	4.97	C
ATOM	1958	O	THR	65B	-43.569	53.289	-69.509	1.00	6.98	O
ATOM	1959	N	TYR	66B	-42.216	52.356	-68.002	1.00	8.08	N
ATOM	1960	CA	TYR	66B	-41.566	51.621	-69.072	1.00	9.06	C
ATOM	1961	CB	TYR	66B	-40.254	51.004	-68.599	1.00	11.12	C
ATOM	1962	CG	TYR	66B	-39.097	51.919	-68.916	1.00	12.07	C
ATOM	1963	CD1	TYR	66B	-38.364	52.533	-67.906	1.00	11.75	C
ATOM	1964	CE1	TYR	66B	-37.360	53.449	-68.210	1.00	13.37	C
ATOM	1965	CD2	TYR	66B	-38.795	52.236	-70.241	1.00	14.31	C
ATOM	1966	CE2	TYR	66B	-37.797	53.150	-70.557	1.00	14.45	C
ATOM	1967	CZ	TYR	66B	-37.087	53.753	-69.540	1.00	15.31	C
ATOM	1968	OH	TYR	66B	-36.116	54.669	-69.859	1.00	15.24	O
ATOM	1969	C	TYR	66B	-42.561	50.560	-69.518	1.00	7.59	C
ATOM	1970	O	TYR	66B	-43.475	50.225	-68.775	1.00	8.33	O
ATOM	1971	N	MET	67B	-42.400	50.060	-70.737	1.00	7.28	N
ATOM	1972	CA	MET	67B	-43.310	49.057	-71.278	1.00	8.78	C
ATOM	1973	CB	MET	67B	-42.746	48.493	-72.589	1.00	7.74	C
ATOM	1974	CG	MET	67B	-43.680	47.533	-73.320	1.00	6.74	C
ATOM	1975	SD	MET	67B	-45.178	48.322	-73.938	1.00	11.73	S
ATOM	1976	CE	MET	67B	-44.710	48.658	-75.652	1.00	9.90	C
ATOM	1977	C	MET	67B	-43.615	47.901	-70.323	1.00	8.91	C
ATOM	1978	O	MET	67B	-44.739	47.411	-70.293	1.00	7.86	O
ATOM	1979	N	ASN	68B	-42.626	47.469	-69.541	1.00	10.61	N
ATOM	1980	CA	ASN	68B	-42.830	46.340	-68.626	1.00	13.25	C
ATOM	1981	CB	ASN	68B	-41.488	45.715	-68.241	1.00	19.28	C
ATOM	1982	CG	ASN	68B	-40.976	44.755	-69.292	1.00	23.82	C
ATOM	1983	OD1	ASN	68B	-40.569	45.167	-70.381	1.00	27.78	O
ATOM	1984	ND2	ASN	68B	-41.009	43.458	-68.978	1.00	25.61	N
ATOM	1985	C	ASN	68B	-43.644	46.596	-67.360	1.00	11.35	C
ATOM	1986	O	ASN	68B	-43.930	45.662	-66.609	1.00	8.74	O
ATOM	1987	N	ARG	69B	-44.017	47.849	-67.123	1.00	9.82	N
ATOM	1988	CA	ARG	69B	-44.818	48.195	-65.956	1.00	8.98	C
ATOM	1989	CB	ARG	69B	-44.010	49.060	-64.989	1.00	10.02	C
ATOM	1990	CG	ARG	69B	-42.903	48.317	-64.265	1.00	11.52	C
ATOM	1991	CD	ARG	69B	-43.470	47.386	-63.206	1.00	11.69	C
ATOM	1992	NE	ARG	69B	-42.418	46.630	-62.537	1.00	15.30	N
ATOM	1993	CZ	ARG	69B	-41.741	45.641	-63.106	1.00	14.17	C
ATOM	1994	NH1	ARG	69B	-40.793	45.006	-62.428	1.00	16.63	N
ATOM	1995	NH2	ARG	69B	-42.026	45.280	-64.348	1.00	14.44	N
ATOM	1996	C	ARG	69B	-46.066	48.954	-66.388	1.00	6.66	C
ATOM	1997	O	ARG	69B	-46.715	49.604	-65.575	1.00	5.49	O
ATOM	1998	N	SER	70B	-46.408	48.859	-67.669	1.00	6.01	N
ATOM	1999	CA	SER	70B	-47.573	49.574	-68.180	1.00	4.00	C
ATOM	2000	CB	SER	70B	-47.634	49.468	-69.709	1.00	4.50	C
ATOM	2001	OG	SER	70B	-47.607	48.129	-70.154	1.00	8.31	O
ATOM	2002	C	SER	70B	-48.884	49.116	-67.545	1.00	3.40	C
ATOM	2003	O	SER	70B	-49.890	49.828	-67.590	1.00	4.53	O
ATOM	2004	N	GLY	71B	-48.864	47.936	-66.936	1.00	3.05	N
ATOM	2005	CA	GLY	71B	-50.049	47.418	-66.280	1.00	3.14	C
ATOM	2006	C	GLY	71B	-50.446	48.255	-65.079	1.00	2.33	C
ATOM	2007	O	GLY	71B	-51.625	48.327	-64.734	1.00	4.33	O
ATOM	2008	N	GLN	72B	-49.470	48.882	-64.425	1.00	1.56	N
ATOM	2009	CA	GLN	72B	-49.778	49.724	-63.276	1.00	4.39	C
ATOM	2010	CB	GLN	72B	-48.493	50.149	-62.556	1.00	7.88	C

FIGURE 9 - 33

```
ATOM   2011  CG  GLN   72B     -47.971  49.091 -61.604  1.00 13.35 C
ATOM   2012  CD  GLN   72B     -46.643  49.454 -60.970  1.00 18.57 C
ATOM   2013  OE1 GLN   72B     -46.510  50.496 -60.323  1.00 22.32 O
ATOM   2014  NE2 GLN   72B     -45.650  48.587 -61.145  1.00 19.02 N
ATOM   2015  C   GLN   72B     -50.550  50.953 -63.745  1.00  2.94 C
ATOM   2016  O   GLN   72B     -51.397  51.478 -63.032  1.00  2.33 O
ATOM   2017  N   SER   73B     -50.245  51.405 -64.955  1.00  3.74 N
ATOM   2018  CA  SER   73B     -50.916  52.558 -65.531  1.00  2.74 C
ATOM   2019  CB  SER   73B     -50.174  53.014 -66.796  1.00  3.50 C
ATOM   2020  OG  SER   73B     -50.846  54.095 -67.422  1.00  2.01 O
ATOM   2021  C   SER   73B     -52.364  52.197 -65.871  1.00  2.96 C
ATOM   2022  O   SER   73B     -53.309  52.861 -65.435  1.00  1.49 O
ATOM   2023  N   VAL   74B     -52.538  51.123 -66.632  1.00  2.23 N
ATOM   2024  CA  VAL   74B     -53.873  50.702 -67.029  1.00  2.54 C
ATOM   2025  CB  VAL   74B     -53.805  49.459 -67.951  1.00  3.62 C
ATOM   2026  CG1 VAL   74B     -53.000  49.790 -69.219  1.00  1.16 C
ATOM   2027  CG2 VAL   74B     -53.183  48.295 -67.217  1.00  9.12 C
ATOM   2028  C   VAL   74B     -54.779  50.407 -65.830  1.00  1.17 C
ATOM   2029  O   VAL   74B     -55.935  50.845 -65.790  1.00  1.00 O
ATOM   2030  N   ALA   75B     -54.249  49.678 -64.851  1.00  1.98 N
ATOM   2031  CA  ALA   75B     -55.007  49.319 -63.654  1.00  2.83 C
ATOM   2032  CB  ALA   75B     -54.200  48.334 -62.804  1.00  3.98 C
ATOM   2033  C   ALA   75B     -55.365  50.558 -62.829  1.00  1.44 C
ATOM   2034  O   ALA   75B     -56.444  50.639 -62.238  1.00  1.56 O
ATOM   2035  N   ALA   76B     -54.465  51.532 -62.792  1.00  3.27 N
ATOM   2036  CA  ALA   76B     -54.720  52.745 -62.034  1.00  1.28 C
ATOM   2037  CB  ALA   76B     -53.464  53.635 -62.023  1.00  1.00 C
ATOM   2038  C   ALA   76B     -55.895  53.520 -62.613  1.00  3.77 C
ATOM   2039  O   ALA   76B     -56.739  54.033 -61.873  1.00  1.48 O
ATOM   2040  N   LEU   77B     -55.946  53.615 -63.937  1.00  3.04 N
ATOM   2041  CA  LEU   77B     -57.019  54.361 -64.583  1.00  3.49 C
ATOM   2042  CB  LEU   77B     -56.587  54.797 -65.985  1.00  4.49 C
ATOM   2043  CG  LEU   77B     -57.535  55.725 -66.759  1.00  4.68 C
ATOM   2044  CD1 LEU   77B     -57.720  57.050 -66.029  1.00  7.73 C
ATOM   2045  CD2 LEU   77B     -56.954  55.961 -68.139  1.00  6.25 C
ATOM   2046  C   LEU   77B     -58.307  53.544 -64.656  1.00  2.77 C
ATOM   2047  O   LEU   77B     -59.385  54.049 -64.362  1.00  2.66 O
ATOM   2048  N   ALA   78B     -58.190  52.279 -65.042  1.00  4.00 N
ATOM   2049  CA  ALA   78B     -59.366  51.420 -65.147  1.00  5.10 C
ATOM   2050  CB  ALA   78B     -58.968  50.067 -65.708  1.00  5.23 C
ATOM   2051  C   ALA   78B     -60.021  51.254 -63.778  1.00  6.08 C
ATOM   2052  O   ALA   78B     -61.247  51.330 -63.645  1.00  6.49 O
ATOM   2053  N   GLY   79B     -59.193  51.045 -62.761  1.00  5.42 N
ATOM   2054  CA  GLY   79B     -59.711  50.865 -61.420  1.00  4.24 C
ATOM   2055  C   GLY   79B     -60.444  52.085 -60.916  1.00  5.54 C
ATOM   2056  O   GLY   79B     -61.544  51.983 -60.365  1.00  3.87 O
ATOM   2057  N   PHE   80B     -59.829  53.246 -61.114  1.00  4.36 N
ATOM   2058  CA  PHE   80B     -60.385  54.515 -60.673  1.00  5.56 C
ATOM   2059  CB  PHE   80B     -59.478  55.655 -61.124  1.00  5.80 C
ATOM   2060  CG  PHE   80B     -59.785  56.960 -60.463  1.00  2.41 C
ATOM   2061  CD1 PHE   80B     -59.175  57.299 -59.261  1.00  2.37 C
ATOM   2062  CD2 PHE   80B     -60.703  57.841 -61.025  1.00  1.22 C
ATOM   2063  CE1 PHE   80B     -59.475  58.505 -58.621  1.00  3.66 C
ATOM   2064  CE2 PHE   80B     -61.012  59.047 -60.394  1.00  4.14 C
ATOM   2065  CZ  PHE   80B     -60.398  59.379 -59.191  1.00  3.11 C
ATOM   2066  C   PHE   80B     -61.798  54.772 -61.191  1.00  5.32 C
ATOM   2067  O   PHE   80B     -62.659  55.280 -60.462  1.00  5.33 O
ATOM   2068  N   PHE   81B     -62.031  54.428 -62.453  1.00  5.93 N
ATOM   2069  CA  PHE   81B     -63.340  54.641 -63.068  1.00  6.93 C
ATOM   2070  CB  PHE   81B     -63.159  55.287 -64.442  1.00  8.01 C
ATOM   2071  CG  PHE   81B     -62.683  56.706 -64.372  1.00  8.00 C
```

FIGURE 9 - 34

```
ATOM   2072  CD1 PHE  81B   -63.537  57.712 -63.921  1.00 10.51 C
ATOM   2073  CD2 PHE  81B   -61.370  57.037 -64.698  1.00  9.92 C
ATOM   2074  CE1 PHE  81B   -63.091  59.022 -63.793  1.00  8.70 C
ATOM   2075  CE2 PHE  81B   -60.912  58.350 -64.572  1.00  9.39 C
ATOM   2076  CZ  PHE  81B   -61.774  59.343 -64.117  1.00  9.39 C
ATOM   2077  C   PHE  81B   -64.171  53.374 -63.187  1.00  8.16 C
ATOM   2078  O   PHE  81B   -65.185  53.353 -63.891  1.00  9.04 O
ATOM   2079  N   ARG  82B   -63.743  52.329 -62.483  1.00  8.06 N
ATOM   2080  CA  ARG  82B   -64.427  51.036 -62.485  1.00  9.95 C
ATOM   2081  CB  ARG  82B   -65.736  51.117 -61.689  1.00 13.35 C
ATOM   2082  CG  ARG  82B   -65.562  51.440 -60.207  1.00 19.57 C
ATOM   2083  CD  ARG  82B   -65.327  52.925 -59.981  1.00 25.43 C
ATOM   2084  NE  ARG  82B   -65.255  53.268 -58.562  1.00 30.97 N
ATOM   2085  CZ  ARG  82B   -64.224  52.991 -57.773  1.00 32.28 C
ATOM   2086  NH1 ARG  82B   -64.254  53.339 -56.495  1.00 33.29 N
ATOM   2087  NH2 ARG  82B   -63.157  52.375 -58.262  1.00 34.75 N
ATOM   2088  C   ARG  82B   -64.716  50.520 -63.893  1.00  8.99 C
ATOM   2089  O   ARG  82B   -65.856  50.175 -64.221  1.00  8.59 O
ATOM   2090  N   ILE  83B   -63.674  50.468 -64.718  1.00  6.16 N
ATOM   2091  CA  ILE  83B   -63.787  49.992 -66.090  1.00  5.00 C
ATOM   2092  CB  ILE  83B   -62.909  50.842 -67.050  1.00  3.62 C
ATOM   2093  CG2 ILE  83B   -62.964  50.265 -68.469  1.00  3.34 C
ATOM   2094  CG1 ILE  83B   -63.384  52.301 -67.034  1.00  4.33 C
ATOM   2095  CD1 ILE  83B   -62.498  53.252 -67.829  1.00  5.71 C
ATOM   2096  C   ILE  83B   -63.353  48.525 -66.167  1.00  4.95 C
ATOM   2097  O   ILE  83B   -62.212  48.181 -65.852  1.00  5.39 O
ATOM   2098  N   ALA  84B   -64.279  47.667 -66.584  1.00  5.42 N
ATOM   2099  CA  ALA  84B   -64.019  46.235 -66.705  1.00  4.82 C
ATOM   2100  CB  ALA  84B   -65.338  45.488 -66.802  1.00  4.52 C
ATOM   2101  C   ALA  84B   -63.145  45.898 -67.913  1.00  5.16 C
ATOM   2102  O   ALA  84B   -63.172  46.591 -68.925  1.00  4.39 O
ATOM   2103  N   PRO  85B   -62.370  44.806 -67.824  1.00  5.84 N
ATOM   2104  CD  PRO  85B   -62.169  43.931 -66.656  1.00  4.89 C
ATOM   2105  CA  PRO  85B   -61.497  44.404 -68.930  1.00  5.59 C
ATOM   2106  CB  PRO  85B   -60.976  43.045 -68.475  1.00  6.29 C
ATOM   2107  CG  PRO  85B   -60.868  43.234 -66.995  1.00  7.89 C
ATOM   2108  C   PRO  85B   -62.217  44.339 -70.272  1.00  5.85 C
ATOM   2109  O   PRO  85B   -61.647  44.705 -71.300  1.00  5.60 O
ATOM   2110  N   ASP  86B   -63.473  43.900 -70.259  1.00  3.97 N
ATOM   2111  CA  ASP  86B   -64.250  43.798 -71.492  1.00  6.06 C
ATOM   2112  CB  ASP  86B   -65.624  43.182 -71.206  1.00  6.85 C
ATOM   2113  CG  ASP  86B   -66.303  42.666 -72.461  1.00 10.85 C
ATOM   2114  OD1 ASP  86B   -65.620  42.009 -73.275  1.00 12.25 O
ATOM   2115  OD2 ASP  86B   -67.521  42.898 -72.632  1.00 11.09 O
ATOM   2116  C   ASP  86B   -64.433  45.145 -72.189  1.00  3.21 C
ATOM   2117  O   ASP  86B   -64.681  45.195 -73.385  1.00  3.89 O
ATOM   2118  N   ALA  87B   -64.303  46.236 -71.443  1.00  3.69 N
ATOM   2119  CA  ALA  87B   -64.468  47.569 -72.015  1.00  2.12 C
ATOM   2120  CB  ALA  87B   -65.267  48.449 -71.055  1.00  2.35 C
ATOM   2121  C   ALA  87B   -63.141  48.246 -72.360  1.00  2.57 C
ATOM   2122  O   ALA  87B   -63.100  49.447 -72.616  1.00  3.65 O
ATOM   2123  N   ILE  88B   -62.065  47.473 -72.368  1.00  3.04 N
ATOM   2124  CA  ILE  88B   -60.746  48.006 -72.679  1.00  2.52 C
ATOM   2125  CB  ILE  88B   -59.727  47.650 -71.568  1.00  2.50 C
ATOM   2126  CG2 ILE  88B   -58.341  48.182 -71.945  1.00  1.08 C
ATOM   2127  CG1 ILE  88B   -60.203  48.207 -70.223  1.00  3.69 C
ATOM   2128  CD1 ILE  88B   -59.321  47.810 -69.035  1.00  2.16 C
ATOM   2129  C   ILE  88B   -60.199  47.482 -74.007  1.00  1.45 C
ATOM   2130  O   ILE  88B   -60.313  46.304 -74.315  1.00  3.14 O
ATOM   2131  N   LEU  89B   -59.608  48.376 -74.789  1.00  3.23 N
ATOM   2132  CA  LEU  89B   -58.989  48.024 -76.066  1.00  1.44 C
```

FIGURE 9 - 35

```
ATOM   2133  CB  LEU   89B    -59.652  48.784 -77.219  1.00  3.11 C
ATOM   2134  CG  LEU   89B    -58.905  48.749 -78.562  1.00  2.58 C
ATOM   2135  CD1 LEU   89B    -58.770  47.305 -79.051  1.00  1.00 C
ATOM   2136  CD2 LEU   89B    -59.660  49.606 -79.587  1.00  1.46 C
ATOM   2137  C   LEU   89B    -57.523  48.441 -75.962  1.00  2.19 C
ATOM   2138  O   LEU   89B    -57.235  49.593 -75.652  1.00  2.78 O
ATOM   2139  N   VAL   90B    -56.600  47.513 -76.204  1.00  1.54 N
ATOM   2140  CA  VAL   90B    -55.180  47.839 -76.124  1.00  1.83 C
ATOM   2141  CB  VAL   90B    -54.452  46.929 -75.107  1.00  4.14 C
ATOM   2142  CG1 VAL   90B    -52.972  47.306 -75.039  1.00  1.73 C
ATOM   2143  CG2 VAL   90B    -55.095  47.057 -73.731  1.00  2.19 C
ATOM   2144  C   VAL   90B    -54.472  47.727 -77.476  1.00  2.14 C
ATOM   2145  O   VAL   90B    -54.399  46.645 -78.062  1.00  1.00 O
ATOM   2146  N   ALA   91B    -53.953  48.849 -77.970  1.00  2.74 N
ATOM   2147  CA  ALA   91B    -53.232  48.864 -79.241  1.00  2.45 C
ATOM   2148  CB  ALA   91B    -53.389  50.221 -79.932  1.00  1.00 C
ATOM   2149  C   ALA   91B    -51.758  48.565 -78.976  1.00  3.66 C
ATOM   2150  O   ALA   91B    -51.150  49.116 -78.058  1.00  2.32 O
ATOM   2151  N   HIS   92B    -51.178  47.695 -79.792  1.00  2.68 N
ATOM   2152  CA  HIS   92B    -49.795  47.318 -79.588  1.00  5.01 C
ATOM   2153  CB  HIS   92B    -49.725  46.243 -78.493  1.00  3.10 C
ATOM   2154  CG  HIS   92B    -50.197  44.893 -78.939  1.00  2.03 C
ATOM   2155  CD2 HIS   92B    -49.538  43.881 -79.548  1.00  1.00 C
ATOM   2156  ND1 HIS   92B    -51.504  44.473 -78.805  1.00  4.66 N
ATOM   2157  CE1 HIS   92B    -51.628  43.257 -79.308  1.00  1.00 C
ATOM   2158  NE2 HIS   92B    -50.448  42.875 -79.766  1.00  1.08 N
ATOM   2159  C   HIS   92B    -49.149  46.797 -80.866  1.00  3.68 C
ATOM   2160  O   HIS   92B    -49.842  46.384 -81.798  1.00  3.85 O
ATOM   2161  N   ASP   93B    -47.820  46.826 -80.902  1.00  4.00 N
ATOM   2162  CA  ASP   93B    -47.067  46.334 -82.050  1.00  1.29 C
ATOM   2163  CB  ASP   93B    -45.605  46.785 -81.960  1.00  2.66 C
ATOM   2164  CG  ASP   93B    -44.962  46.446 -80.621  1.00  1.00 C
ATOM   2165  OD1 ASP   93B    -43.926  45.745 -80.607  1.00  3.29 O
ATOM   2166  OD2 ASP   93B    -45.485  46.888 -79.582  1.00  2.95 O
ATOM   2167  C   ASP   93B    -47.159  44.812 -82.096  1.00  2.33 C
ATOM   2168  O   ASP   93B    -47.123  44.147 -81.060  1.00  2.27 O
ATOM   2169  N   GLU   94B    -47.301  44.266 -83.299  1.00  2.49 N
ATOM   2170  CA  GLU   94B    -47.419  42.822 -83.479  1.00  1.67 C
ATOM   2171  CB  GLU   94B    -48.801  42.482 -84.040  1.00  1.00 C
ATOM   2172  CG  GLU   94B    -49.007  40.996 -84.320  1.00  1.84 C
ATOM   2173  CD  GLU   94B    -48.836  40.156 -83.072  1.00  4.62 C
ATOM   2174  OE1 GLU   94B    -49.629  40.323 -82.123  1.00  3.08 O
ATOM   2175  OE2 GLU   94B    -47.897  39.336 -83.039  1.00  9.18 O
ATOM   2176  C   GLU   94B    -46.342  42.268 -84.409  1.00  1.13 C
ATOM   2177  O   GLU   94B    -46.314  42.582 -85.598  1.00  1.00 O
ATOM   2178  N   LEU   95B    -45.481  41.418 -83.861  1.00  1.67 N
ATOM   2179  CA  LEU   95B    -44.387  40.813 -84.619  1.00  1.00 C
ATOM   2180  CB  LEU   95B    -43.424  40.096 -83.657  1.00  2.42 C
ATOM   2181  CG  LEU   95B    -42.610  40.972 -82.688  1.00  1.60 C
ATOM   2182  CD1 LEU   95B    -42.106  40.143 -81.518  1.00  1.61 C
ATOM   2183  CD2 LEU   95B    -41.456  41.619 -83.428  1.00  2.70 C
ATOM   2184  C   LEU   95B    -44.850  39.816 -85.680  1.00  2.71 C
ATOM   2185  O   LEU   95B    -44.208  39.664 -86.723  1.00  1.95 O
ATOM   2186  N   ASP   96B    -45.968  39.148 -85.423  1.00  3.16 N
ATOM   2187  CA  ASP   96B    -46.460  38.125 -86.337  1.00  2.26 C
ATOM   2188  CB  ASP   96B    -47.302  37.108 -85.568  1.00  4.92 C
ATOM   2189  CG  ASP   96B    -46.463  36.239 -84.650  1.00  4.83 C
ATOM   2190  OD1 ASP   96B    -45.223  36.197 -84.824  1.00  6.60 O
ATOM   2191  OD2 ASP   96B    -47.048  35.583 -83.763  1.00  8.23 O
ATOM   2192  C   ASP   96B    -47.209  38.575 -87.579  1.00  4.06 C
ATOM   2193  O   ASP   96B    -47.672  37.741 -88.359  1.00  1.38 O
```

FIGURE 9 - 36

ATOM	2194	N	MET	97B	-47.324	39.887	-87.758	1.00	1.92	N
ATOM	2195	CA	MET	97B	-47.982	40.464	-88.922	1.00	4.52	C
ATOM	2196	CB	MET	97B	-49.296	41.126	-88.521	1.00	6.05	C
ATOM	2197	CG	MET	97B	-50.325	40.148	-88.014	1.00	5.91	C
ATOM	2198	SD	MET	97B	-51.790	40.977	-87.382	1.00	10.42	S
ATOM	2199	CE	MET	97B	-52.329	39.813	-86.176	1.00	6.99	C
ATOM	2200	C	MET	97B	-47.033	41.495	-89.514	1.00	4.54	C
ATOM	2201	O	MET	97B	-46.381	42.234	-88.785	1.00	2.81	O
ATOM	2202	N	PRO	98B	-46.933	41.544	-90.847	1.00	4.20	N
ATOM	2203	CD	PRO	98B	-47.684	40.737	-91.831	1.00	6.83	C
ATOM	2204	CA	PRO	98B	-46.048	42.496	-91.520	1.00	5.03	C
ATOM	2205	CB	PRO	98B	-46.024	41.974	-92.949	1.00	5.36	C
ATOM	2206	CG	PRO	98B	-47.440	41.492	-93.126	1.00	6.31	C
ATOM	2207	C	PRO	98B	-46.559	43.929	-91.462	1.00	3.83	C
ATOM	2208	O	PRO	98B	-47.741	44.169	-91.212	1.00	4.97	O
ATOM	2209	N	PRO	99B	-45.666	44.904	-91.679	1.00	4.29	N
ATOM	2210	CD	PRO	99B	-44.201	44.809	-91.774	1.00	1.78	C
ATOM	2211	CA	PRO	99B	-46.102	46.301	-91.651	1.00	3.53	C
ATOM	2212	CB	PRO	99B	-44.835	47.058	-92.036	1.00	3.55	C
ATOM	2213	CG	PRO	99B	-43.763	46.228	-91.427	1.00	2.70	C
ATOM	2214	C	PRO	99B	-47.212	46.463	-92.692	1.00	4.96	C
ATOM	2215	O	PRO	99B	-47.126	45.888	-93.774	1.00	5.42	O
ATOM	2216	N	GLY	100B	-48.255	47.222	-92.362	1.00	6.04	N
ATOM	2217	CA	GLY	100B	-49.345	47.416	-93.304	1.00	4.44	C
ATOM	2218	C	GLY	100B	-50.570	46.577	-92.986	1.00	6.31	C
ATOM	2219	O	GLY	100B	-51.618	46.704	-93.622	1.00	4.97	O
ATOM	2220	N	VAL	101B	-50.446	45.715	-91.988	1.00	5.55	N
ATOM	2221	CA	VAL	101B	-51.553	44.856	-91.599	1.00	8.08	C
ATOM	2222	CB	VAL	101B	-51.201	43.370	-91.837	1.00	8.84	C
ATOM	2223	CG1	VAL	101B	-52.310	42.465	-91.306	1.00	8.12	C
ATOM	2224	CG2	VAL	101B	-50.993	43.135	-93.325	1.00	7.19	C
ATOM	2225	C	VAL	101B	-51.940	45.060	-90.140	1.00	10.30	C
ATOM	2226	O	VAL	101B	-51.079	45.098	-89.256	1.00	10.96	O
ATOM	2227	N	ALA	102B	-53.244	45.194	-89.904	1.00	9.40	N
ATOM	2228	CA	ALA	102B	-53.772	45.390	-88.564	1.00	8.25	C
ATOM	2229	CB	ALA	102B	-54.176	46.852	-88.362	1.00	8.84	C
ATOM	2230	C	ALA	102B	-54.972	44.484	-88.335	1.00	7.82	C
ATOM	2231	O	ALA	102B	-55.860	44.381	-89.182	1.00	6.14	O
ATOM	2232	N	LYS	103B	-54.990	43.820	-87.187	1.00	7.51	N
ATOM	2233	CA	LYS	103B	-56.095	42.940	-86.854	1.00	8.21	C
ATOM	2234	CB	LYS	103B	-55.711	41.480	-87.112	1.00	11.34	C
ATOM	2235	CG	LYS	103B	-55.466	41.168	-88.577	1.00	14.53	C
ATOM	2236	CD	LYS	103B	-55.196	39.687	-88.805	1.00	18.62	C
ATOM	2237	CE	LYS	103B	-54.758	39.424	-90.243	1.00	20.77	C
ATOM	2238	NZ	LYS	103B	-55.747	39.935	-91.238	1.00	25.06	N
ATOM	2239	C	LYS	103B	-56.517	43.133	-85.404	1.00	7.70	C
ATOM	2240	O	LYS	103B	-55.767	43.678	-84.590	1.00	5.76	O
ATOM	2241	N	LEU	104B	-57.730	42.691	-85.095	1.00	6.01	N
ATOM	2242	CA	LEU	104B	-58.282	42.809	-83.753	1.00	5.64	C
ATOM	2243	CB	LEU	104B	-59.658	43.478	-83.810	1.00	4.96	C
ATOM	2244	CG	LEU	104B	-59.657	44.971	-84.135	1.00	5.33	C
ATOM	2245	CD1	LEU	104B	-61.024	45.424	-84.615	1.00	3.79	C
ATOM	2246	CD2	LEU	104B	-59.242	45.726	-82.900	1.00	3.42	C
ATOM	2247	C	LEU	104B	-58.408	41.428	-83.130	1.00	5.65	C
ATOM	2248	O	LEU	104B	-58.615	40.439	-83.832	1.00	5.29	O
ATOM	2249	N	LYS	105B	-58.282	41.364	-81.811	1.00	5.44	N
ATOM	2250	CA	LYS	105B	-58.384	40.091	-81.113	1.00	9.90	C
ATOM	2251	CB	LYS	105B	-57.018	39.400	-81.069	1.00	9.41	C
ATOM	2252	CG	LYS	105B	-57.028	38.090	-80.307	1.00	13.46	C
ATOM	2253	CD	LYS	105B	-55.673	37.420	-80.315	1.00	17.56	C
ATOM	2254	CE	LYS	105B	-55.770	36.030	-79.691	1.00	19.05	C

FIGURE 9 - 37

```
ATOM   2255  NZ  LYS  105B    -54.451  35.356 -79.613  1.00 21.08 N
ATOM   2256  C   LYS  105B    -58.897  40.276 -79.696  1.00 10.24 C
ATOM   2257  O   LYS  105B    -58.593  41.268 -79.040  1.00  8.95 O
ATOM   2258  N   THR  106B    -59.675  39.310 -79.226  1.00 12.32 N
ATOM   2259  CA  THR  106B    -60.206  39.369 -77.878  1.00 15.54 C
ATOM   2260  CB  THR  106B    -61.746  39.250 -77.883  1.00 19.09 C
ATOM   2261  OG1 THR  106B    -62.222  39.131 -76.536  1.00 22.18 O
ATOM   2262  CG2 THR  106B    -62.189  38.043 -78.695  1.00 21.39 C
ATOM   2263  C   THR  106B    -59.587  38.242 -77.053  1.00 15.37 C
ATOM   2264  O   THR  106B    -59.593  37.081 -77.465  1.00 13.68 O
ATOM   2265  N   GLY  107B    -59.026  38.599 -75.903  1.00 14.38 N
ATOM   2266  CA  GLY  107B    -58.416  37.608 -75.039  1.00 15.07 C
ATOM   2267  C   GLY  107B    -57.126  37.032 -75.587  1.00 15.75 C
ATOM   2268  O   GLY  107B    -56.468  37.640 -76.432  1.00 14.52 O
ATOM   2269  N   GLY  108B    -56.766  35.848 -75.106  1.00 16.62 N
ATOM   2270  CA  GLY  108B    -55.542  35.210 -75.557  1.00 18.82 C
ATOM   2271  C   GLY  108B    -54.374  35.569 -74.658  1.00 19.50 C
ATOM   2272  O   GLY  108B    -54.553  36.256 -73.652  1.00 21.15 O
ATOM   2273  N   GLY  109B    -53.180  35.109 -75.021  1.00 18.99 N
ATOM   2274  CA  GLY  109B    -51.997  35.393 -74.225  1.00 18.51 C
ATOM   2275  C   GLY  109B    -51.373  36.738 -74.546  1.00 17.94 C
ATOM   2276  O   GLY  109B    -52.056  37.649 -75.006  1.00 17.28 O
ATOM   2277  N   HIS  110B    -50.073  36.866 -74.300  1.00 17.07 N
ATOM   2278  CA  HIS  110B    -49.371  38.110 -74.572  1.00 16.68 C
ATOM   2279  CB  HIS  110B    -48.585  38.561 -73.338  1.00 15.45 C
ATOM   2280  CG  HIS  110B    -47.773  37.474 -72.709  1.00 13.99 C
ATOM   2281  CD2 HIS  110B    -46.544  36.992 -73.009  1.00 12.22 C
ATOM   2282  ND1 HIS  110B    -48.226  36.730 -71.642  1.00 14.67 N
ATOM   2283  CE1 HIS  110B    -47.310  35.837 -71.311  1.00 12.51 C
ATOM   2284  NE2 HIS  110B    -46.280  35.975 -72.125  1.00 13.04 N
ATOM   2285  C   HIS  110B    -48.421  37.976 -75.754  1.00 17.20 C
ATOM   2286  O   HIS  110B    -47.856  38.967 -76.219  1.00 15.23 O
ATOM   2287  N   GLY  111B    -48.247  36.748 -76.235  1.00 17.56 N
ATOM   2288  CA  GLY  111B    -47.355  36.518 -77.359  1.00 18.79 C
ATOM   2289  C   GLY  111B    -45.929  36.943 -77.066  1.00 18.68 C
ATOM   2290  O   GLY  111B    -45.158  37.243 -77.980  1.00 19.46 O
ATOM   2291  N   GLY  112B    -45.575  36.965 -75.785  1.00 18.29 N
ATOM   2292  CA  GLY  112B    -44.237  37.362 -75.391  1.00 16.30 C
ATOM   2293  C   GLY  112B    -44.087  38.870 -75.298  1.00 15.47 C
ATOM   2294  O   GLY  112B    -43.011  39.374 -74.995  1.00 16.30 O
ATOM   2295  N   HIS  113B    -45.170  39.598 -75.555  1.00 13.77 N
ATOM   2296  CA  HIS  113B    -45.124  41.052 -75.493  1.00 11.72 C
ATOM   2297  CB  HIS  113B    -46.330  41.649 -76.227  1.00  9.23 C
ATOM   2298  CG  HIS  113B    -46.215  43.118 -76.482  1.00  5.60 C
ATOM   2299  CD2 HIS  113B    -45.995  43.811 -77.625  1.00  5.62 C
ATOM   2300  ND1 HIS  113B    -46.327  44.058 -75.481  1.00  5.15 N
ATOM   2301  CE1 HIS  113B    -46.186  45.266 -75.996  1.00  4.67 C
ATOM   2302  NE2 HIS  113B    -45.983  45.144 -77.295  1.00  5.20 N
ATOM   2303  C   HIS  113B    -45.105  41.499 -74.031  1.00 11.35 C
ATOM   2304  O   HIS  113B    -46.061  41.283 -73.291  1.00  9.97 O
ATOM   2305  N   ASN  114B    -43.992  42.107 -73.635  1.00 11.65 N
ATOM   2306  CA  ASN  114B    -43.777  42.594 -72.277  1.00 13.51 C
ATOM   2307  CB  ASN  114B    -42.422  43.300 -72.209  1.00 20.27 C
ATOM   2308  CG  ASN  114B    -41.958  43.794 -73.570  1.00 25.59 C
ATOM   2309  OD1 ASN  114B    -41.706  42.998 -74.483  1.00 29.37 O
ATOM   2310  ND2 ASN  114B    -41.847  45.107 -73.716  1.00 26.71 N
ATOM   2311  C   ASN  114B    -44.874  43.519 -71.752  1.00 11.42 C
ATOM   2312  O   ASN  114B    -45.242  43.449 -70.576  1.00  6.04 O
ATOM   2313  N   GLY  115B    -45.390  44.384 -72.623  1.00  8.20 N
ATOM   2314  CA  GLY  115B    -46.445  45.295 -72.218  1.00  6.57 C
ATOM   2315  C   GLY  115B    -47.737  44.546 -71.948  1.00  4.39 C
```

FIGURE 9 - 38

```
ATOM   2316  O   GLY  115B    -48.354  44.707 -70.895  1.00  4.47 O
ATOM   2317  N   LEU  116B    -48.154  43.722 -72.901  1.00  5.66 N
ATOM   2318  CA  LEU  116B    -49.377  42.946 -72.733  1.00  5.91 C
ATOM   2319  CB  LEU  116B    -49.658  42.108 -73.982  1.00  9.34 C
ATOM   2320  CG  LEU  116B    -50.017  42.843 -75.276  1.00 12.19 C
ATOM   2321  CD1 LEU  116B    -50.256  41.820 -76.381  1.00 13.60 C
ATOM   2322  CD2 LEU  116B    -51.258  43.707 -75.064  1.00 12.16 C
ATOM   2323  C   LEU  116B    -49.247  42.027 -71.517  1.00  5.65 C
ATOM   2324  O   LEU  116B    -50.198  41.838 -70.759  1.00  4.54 O
ATOM   2325  N   LYS  117B    -48.061  41.462 -71.330  1.00  4.02 N
ATOM   2326  CA  LYS  117B    -47.831  40.574 -70.201  1.00  4.05 C
ATOM   2327  CB  LYS  117B    -46.386  40.071 -70.202  1.00  4.86 C
ATOM   2328  CG  LYS  117B    -46.053  39.194 -69.006  1.00  8.47 C
ATOM   2329  CD  LYS  117B    -44.718  38.488 -69.188  1.00  7.05 C
ATOM   2330  CE  LYS  117B    -43.566  39.463 -69.370  1.00  8.62 C
ATOM   2331  NZ  LYS  117B    -42.313  38.741 -69.740  1.00 13.80 N
ATOM   2332  C   LYS  117B    -48.131  41.261 -68.876  1.00  5.02 C
ATOM   2333  O   LYS  117B    -48.867  40.721 -68.041  1.00  5.03 O
ATOM   2334  N   ASP  118B    -47.587  42.465 -68.694  1.00  3.29 N
ATOM   2335  CA  ASP  118B    -47.798  43.196 -67.455  1.00  3.14 C
ATOM   2336  CB  ASP  118B    -46.904  44.433 -67.387  1.00  2.01 C
ATOM   2337  CG  ASP  118B    -46.802  44.986 -65.980  1.00  4.42 C
ATOM   2338  OD1 ASP  118B    -46.351  44.237 -65.086  1.00  3.69 O
ATOM   2339  OD2 ASP  118B    -47.177  46.154 -65.756  1.00  3.76 O
ATOM   2340  C   ASP  118B    -49.250  43.620 -67.283  1.00  2.90 C
ATOM   2341  O   ASP  118B    -49.749  43.701 -66.161  1.00  3.26 O
ATOM   2342  N   ILE  119B    -49.925  43.909 -68.390  1.00  2.60 N
ATOM   2343  CA  ILE  119B    -51.318  44.307 -68.300  1.00  2.33 C
ATOM   2344  CB  ILE  119B    -51.832  44.853 -69.645  1.00  1.81 C
ATOM   2345  CG2 ILE  119B    -53.340  45.001 -69.596  1.00  1.45 C
ATOM   2346  CG1 ILE  119B    -51.148  46.195 -69.940  1.00  1.00 C
ATOM   2347  CD1 ILE  119B    -51.487  46.805 -71.295  1.00  1.90 C
ATOM   2348  C   ILE  119B    -52.153  43.114 -67.845  1.00  3.28 C
ATOM   2349  O   ILE  119B    -52.981  43.235 -66.945  1.00  1.53 O
ATOM   2350  N   ILE  120B    -51.922  41.960 -68.457  1.00  4.47 N
ATOM   2351  CA  ILE  120B    -52.650  40.754 -68.073  1.00  7.07 C
ATOM   2352  CB  ILE  120B    -52.214  39.542 -68.923  1.00  6.84 C
ATOM   2353  CG2 ILE  120B    -52.792  38.255 -68.338  1.00  9.94 C
ATOM   2354  CG1 ILE  120B    -52.672  39.730 -70.369  1.00  8.07 C
ATOM   2355  CD1 ILE  120B    -52.164  38.659 -71.322  1.00  8.00 C
ATOM   2356  C   ILE  120B    -52.383  40.456 -66.599  1.00  8.10 C
ATOM   2357  O   ILE  120B    -53.295  40.109 -65.850  1.00 10.68 O
ATOM   2358  N   ALA  121B    -51.131  40.611 -66.183  1.00 10.38 N
ATOM   2359  CA  ALA  121B    -50.753  40.353 -64.795  1.00 10.80 C
ATOM   2360  CB  ALA  121B    -49.242  40.424 -64.651  1.00  8.94 C
ATOM   2361  C   ALA  121B    -51.412  41.325 -63.820  1.00 11.72 C
ATOM   2362  O   ALA  121B    -51.958  40.916 -62.796  1.00 12.69 O
ATOM   2363  N   GLN  122B    -51.373  42.611 -64.149  1.00 13.51 N
ATOM   2364  CA  GLN  122B    -51.934  43.644 -63.288  1.00 17.15 C
ATOM   2365  CB  GLN  122B    -51.478  45.020 -63.780  1.00 19.73 C
ATOM   2366  CG  GLN  122B    -51.375  46.081 -62.696  1.00 24.59 C
ATOM   2367  CD  GLN  122B    -50.388  45.711 -61.602  1.00 27.10 C
ATOM   2368  OE1 GLN  122B    -50.634  44.804 -60.808  1.00 27.53 O
ATOM   2369  NE2 GLN  122B    -49.261  46.415 -61.558  1.00 27.90 N
ATOM   2370  C   GLN  122B    -53.458  43.616 -63.168  1.00 18.50 C
ATOM   2371  O   GLN  122B    -54.007  44.025 -62.145  1.00 19.68 O
ATOM   2372  N   LEU  123B    -54.143  43.140 -64.203  1.00 17.95 N
ATOM   2373  CA  LEU  123B    -55.601  43.087 -64.170  1.00 21.05 C
ATOM   2374  CB  LEU  123B    -56.171  43.214 -65.586  1.00 16.96 C
ATOM   2375  CG  LEU  123B    -55.964  44.576 -66.256  1.00 17.66 C
ATOM   2376  CD1 LEU  123B    -56.508  44.543 -67.679  1.00 12.24 C
```

FIGURE 9 - 39

```
ATOM   2377  CD2 LEU   123B    -56.657  45.656 -65.439  1.00 15.19 C
ATOM   2378  C   LEU   123B    -56.112  41.808 -63.515  1.00 24.78 C
ATOM   2379  O   LEU   123B    -57.198  41.324 -63.838  1.00 24.24 O
ATOM   2380  N   GLY   124B    -55.327  41.273 -62.585  1.00 27.91 N
ATOM   2381  CA  GLY   124B    -55.713  40.057 -61.893  1.00 32.42 C
ATOM   2382  C   GLY   124B    -55.957  38.883 -62.822  1.00 35.00 C
ATOM   2383  O   GLY   124B    -57.071  38.361 -62.892  1.00 36.71 O
ATOM   2384  N   ASN   125B    -54.915  38.467 -63.536  1.00 37.45 N
ATOM   2385  CA  ASN   125B    -55.003  37.342 -64.466  1.00 38.98 C
ATOM   2386  CB  ASN   125B    -55.249  36.039 -63.690  1.00 43.03 C
ATOM   2387  CG  ASN   125B    -55.027  34.791 -64.540  1.00 46.70 C
ATOM   2388  OD1 ASN   125B    -55.800  34.494 -65.452  1.00 48.47 O
ATOM   2389  ND2 ASN   125B    -53.959  34.056 -64.239  1.00 48.29 N
ATOM   2390  C   ASN   125B    -56.104  37.552 -65.509  1.00 37.04 C
ATOM   2391  O   ASN   125B    -56.509  36.613 -66.195  1.00 37.49 O
ATOM   2392  N   ASN   126B    -56.584  38.787 -65.628  1.00 35.87 N
ATOM   2393  CA  ASN   126B    -57.628  39.098 -66.599  1.00 35.01 C
ATOM   2394  CB  ASN   126B    -58.524  40.233 -66.093  1.00 34.84 C
ATOM   2395  CG  ASN   126B    -59.512  39.773 -65.038  1.00 34.58 C
ATOM   2396  OD1 ASN   126B    -60.270  40.574 -64.491  1.00 35.41 O
ATOM   2397  ND2 ASN   126B    -59.514  38.476 -64.751  1.00 35.15 N
ATOM   2398  C   ASN   126B    -57.036  39.483 -67.947  1.00 34.96 C
ATOM   2399  O   ASN   126B    -56.297  40.460 -68.059  1.00 33.64 O
ATOM   2400  N   ASN   127B    -57.370  38.703 -68.968  1.00 33.78 N
ATOM   2401  CA  ASN   127B    -56.882  38.953 -70.315  1.00 33.02 C
ATOM   2402  CB  ASN   127B    -55.964  37.810 -70.764  1.00 35.98 C
ATOM   2403  CG  ASN   127B    -56.668  36.462 -70.786  1.00 37.66 C
ATOM   2404  OD1 ASN   127B    -56.054  35.435 -71.079  1.00 39.57 O
ATOM   2405  ND2 ASN   127B    -57.959  36.458 -70.478  1.00 39.63 N
ATOM   2406  C   ASN   127B    -58.062  39.079 -71.268  1.00 29.95 C
ATOM   2407  O   ASN   127B    -57.891  39.100 -72.487  1.00 31.65 O
ATOM   2408  N   SER   128B    -59.260  39.168 -70.698  1.00 27.17 N
ATOM   2409  CA  SER   128B    -60.478  39.276 -71.488  1.00 22.22 C
ATOM   2410  CB  SER   128B    -61.699  38.919 -70.635  1.00 24.63 C
ATOM   2411  OG  SER   128B    -61.941  39.904 -69.645  1.00 27.45 O
ATOM   2412  C   SER   128B    -60.662  40.666 -72.086  1.00 17.26 C
ATOM   2413  O   SER   128B    -61.791  41.102 -72.316  1.00 15.41 O
ATOM   2414  N   PHE   129B    -59.557  41.366 -72.330  1.00 12.07 N
ATOM   2415  CA  PHE   129B    -59.642  42.689 -72.934  1.00  8.40 C
ATOM   2416  CB  PHE   129B    -58.677  43.671 -72.259  1.00  5.02 C
ATOM   2417  CG  PHE   129B    -57.243  43.231 -72.271  1.00  4.53 C
ATOM   2418  CD1 PHE   129B    -56.397  43.587 -73.314  1.00  3.53 C
ATOM   2419  CD2 PHE   129B    -56.731  42.480 -71.220  1.00  3.19 C
ATOM   2420  CE1 PHE   129B    -55.054  43.206 -73.303  1.00  3.46 C
ATOM   2421  CE2 PHE   129B    -55.391  42.093 -71.200  1.00  4.90 C
ATOM   2422  CZ  PHE   129B    -54.553  42.459 -72.241  1.00  3.02 C
ATOM   2423  C   PHE   129B    -59.339  42.555 -74.418  1.00  6.14 C
ATOM   2424  O   PHE   129B    -58.816  41.539 -74.862  1.00  6.28 O
ATOM   2425  N   HIS   130B    -59.672  43.577 -75.192  1.00  4.10 N
ATOM   2426  CA  HIS   130B    -59.458  43.513 -76.622  1.00  3.45 C
ATOM   2427  CB  HIS   130B    -60.638  44.179 -77.327  1.00  5.34 C
ATOM   2428  CG  HIS   130B    -61.962  43.647 -76.876  1.00  8.98 C
ATOM   2429  CD2 HIS   130B    -62.657  42.552 -77.263  1.00  8.24 C
ATOM   2430  ND1 HIS   130B    -62.690  44.222 -75.855  1.00 10.64 N
ATOM   2431  CE1 HIS   130B    -63.776  43.503 -75.634  1.00  9.87 C
ATOM   2432  NE2 HIS   130B    -63.780  42.484 -76.474  1.00 12.29 N
ATOM   2433  C   HIS   130B    -58.127  44.108 -77.053  1.00  3.41 C
ATOM   2434  O   HIS   130B    -57.550  44.938 -76.350  1.00  1.40 O
ATOM   2435  N   ARG   131B    -57.627  43.651 -78.196  1.00  3.01 N
ATOM   2436  CA  ARG   131B    -56.350  44.131 -78.698  1.00  2.63 C
ATOM   2437  CB  ARG   131B    -55.258  43.066 -78.508  1.00  5.78 C
```

FIGURE 9 - 40

```
ATOM   2438  CG  ARG   131B    -55.131  42.487 -77.109  1.00  7.17 C
ATOM   2439  CD  ARG   131B    -56.097  41.321 -76.911  1.00  9.66 C
ATOM   2440  NE  ARG   131B    -55.919  40.650 -75.626  1.00 10.93 N
ATOM   2441  CZ  ARG   131B    -54.884  39.874 -75.314  1.00 13.28 C
ATOM   2442  NH1 ARG   131B    -53.914  39.660 -76.194  1.00 11.77 N
ATOM   2443  NH2 ARG   131B    -54.826  39.295 -74.119  1.00 13.11 N
ATOM   2444  C   ARG   131B    -56.356  44.524 -80.170  1.00  4.24 C
ATOM   2445  O   ARG   131B    -56.925  43.815 -81.011  1.00  2.91 O
ATOM   2446  N   LEU   132B    -55.730  45.664 -80.465  1.00  3.65 N
ATOM   2447  CA  LEU   132B    -55.561  46.127 -81.837  1.00  1.00 C
ATOM   2448  CB  LEU   132B    -55.796  47.643 -81.971  1.00  1.80 C
ATOM   2449  CG  LEU   132B    -55.415  48.216 -83.348  1.00  1.00 C
ATOM   2450  CD1 LEU   132B    -56.264  47.555 -84.426  1.00  1.28 C
ATOM   2451  CD2 LEU   132B    -55.606  49.750 -83.370  1.00  1.00 C
ATOM   2452  C   LEU   132B    -54.098  45.787 -82.132  1.00  2.48 C
ATOM   2453  O   LEU   132B    -53.171  46.391 -81.581  1.00  1.00 O
ATOM   2454  N   ARG   133B    -53.898  44.785 -82.978  1.00  1.00 N
ATOM   2455  CA  ARG   133B    -52.564  44.335 -83.338  1.00  1.82 C
ATOM   2456  CB  ARG   133B    -52.592  42.824 -83.576  1.00  2.80 C
ATOM   2457  CG  ARG   133B    -53.305  42.062 -82.457  1.00  2.92 C
ATOM   2458  CD  ARG   133B    -53.561  40.608 -82.824  1.00  5.02 C
ATOM   2459  NE  ARG   133B    -52.331  39.828 -82.904  1.00  6.40 N
ATOM   2460  CZ  ARG   133B    -52.283  38.552 -83.262  1.00 10.22 C
ATOM   2461  NH1 ARG   133B    -53.398  37.906 -83.578  1.00 10.19 N
ATOM   2462  NH2 ARG   133B    -51.116  37.917 -83.300  1.00 12.60 N
ATOM   2463  C   ARG   133B    -52.073  45.061 -84.582  1.00  2.03 C
ATOM   2464  O   ARG   133B    -52.627  44.899 -85.667  1.00  3.66 O
ATOM   2465  N   LEU   134B    -51.033  45.869 -84.413  1.00  1.00 N
ATOM   2466  CA  LEU   134B    -50.463  46.625 -85.522  1.00  2.19 C
ATOM   2467  CB  LEU   134B    -50.168  48.067 -85.094  1.00  2.72 C
ATOM   2468  CG  LEU   134B    -51.342  48.904 -84.575  1.00  5.54 C
ATOM   2469  CD1 LEU   134B    -52.394  49.088 -85.666  1.00  3.29 C
ATOM   2470  CD2 LEU   134B    -51.942  48.218 -83.373  1.00  4.70 C
ATOM   2471  C   LEU   134B    -49.183  45.942 -85.965  1.00  2.04 C
ATOM   2472  O   LEU   134B    -48.174  45.963 -85.253  1.00  2.64 O
ATOM   2473  N   GLY   135B    -49.231  45.325 -87.139  1.00  2.59 N
ATOM   2474  CA  GLY   135B    -48.065  44.632 -87.656  1.00  2.15 C
ATOM   2475  C   GLY   135B    -46.815  45.484 -87.791  1.00  2.34 C
ATOM   2476  O   GLY   135B    -46.875  46.633 -88.219  1.00  5.39 O
ATOM   2477  N   ILE   136B    -45.674  44.914 -87.413  1.00  2.33 N
ATOM   2478  CA  ILE   136B    -44.391  45.602 -87.520  1.00  2.37 C
ATOM   2479  CB  ILE   136B    -43.865  46.092 -86.151  1.00  2.52 C
ATOM   2480  CG2 ILE   136B    -44.721  47.236 -85.639  1.00  1.16 C
ATOM   2481  CG1 ILE   136B    -43.827  44.932 -85.156  1.00  1.83 C
ATOM   2482  CD1 ILE   136B    -42.898  45.174 -83.982  1.00  4.09 C
ATOM   2483  C   ILE   136B    -43.361  44.636 -88.092  1.00  2.50 C
ATOM   2484  O   ILE   136B    -42.197  44.984 -88.253  1.00  2.54 O
ATOM   2485  N   GLY   137B    -43.799  43.418 -88.391  1.00  4.78 N
ATOM   2486  CA  GLY   137B    -42.889  42.425 -88.935  1.00  2.44 C
ATOM   2487  C   GLY   137B    -42.043  41.746 -87.870  1.00  1.61 C
ATOM   2488  O   GLY   137B    -42.038  42.158 -86.704  1.00  1.06 O
ATOM   2489  N   HIS   138B    -41.323  40.706 -88.281  1.00  1.34 N
ATOM   2490  CA  HIS   138B    -40.461  39.935 -87.384  1.00  1.00 C
ATOM   2491  CB  HIS   138B    -41.046  38.534 -87.168  1.00  1.67 C
ATOM   2492  CG  HIS   138B    -40.357  37.745 -86.096  1.00  1.76 C
ATOM   2493  CD2 HIS   138B    -39.104  37.238 -86.027  1.00  3.34 C
ATOM   2494  ND1 HIS   138B    -40.962  37.436 -84.896  1.00  4.71 N
ATOM   2495  CE1 HIS   138B    -40.108  36.779 -84.131  1.00  3.61 C
ATOM   2496  NE2 HIS   138B    -38.973  36.646 -84.794  1.00  3.60 N
ATOM   2497  C   HIS   138B    -39.071  39.813 -88.014  1.00  3.36 C
ATOM   2498  O   HIS   138B    -38.950  39.554 -89.211  1.00  6.22 O
```

FIGURE 9 - 41

```
ATOM   2499  N   PRO  139B    -38.007  39.993 -87.213  1.00  4.22 N
ATOM   2500  CD  PRO  139B    -38.061  40.396 -85.796  1.00  4.12 C
ATOM   2501  CA  PRO  139B    -36.613  39.911 -87.669  1.00  4.56 C
ATOM   2502  CB  PRO  139B    -35.837  40.481 -86.484  1.00  4.77 C
ATOM   2503  CG  PRO  139B    -36.671  40.053 -85.318  1.00  3.14 C
ATOM   2504  C   PRO  139B    -36.114  38.521 -88.083  1.00  6.85 C
ATOM   2505  O   PRO  139B    -35.112  38.404 -88.785  1.00  5.91 O
ATOM   2506  N   GLY  140B    -36.802  37.472 -87.653  1.00  8.55 N
ATOM   2507  CA  GLY  140B    -36.384  36.130 -88.016  1.00 11.48 C
ATOM   2508  C   GLY  140B    -36.188  35.205 -86.831  1.00 13.15 C
ATOM   2509  O   GLY  140B    -36.931  34.235 -86.661  1.00 16.89 O
ATOM   2510  N   HIS  141B    -35.193  35.497 -86.003  1.00 14.08 N
ATOM   2511  CA  HIS  141B    -34.922  34.669 -84.837  1.00 12.93 C
ATOM   2512  CB  HIS  141B    -33.430  34.339 -84.758  1.00 13.74 C
ATOM   2513  CG  HIS  141B    -33.112  33.252 -83.781  1.00 13.95 C
ATOM   2514  CD2 HIS  141B    -32.589  33.300 -82.535  1.00 15.78 C
ATOM   2515  ND1 HIS  141B    -33.394  31.926 -84.028  1.00 13.89 N
ATOM   2516  CE1 HIS  141B    -33.060  31.204 -82.974  1.00 14.40 C
ATOM   2517  NE2 HIS  141B    -32.569  32.014 -82.053  1.00 16.37 N
ATOM   2518  C   HIS  141B    -35.360  35.381 -83.558  1.00 13.42 C
ATOM   2519  O   HIS  141B    -35.219  36.596 -83.436  1.00 14.28 O
ATOM   2520  N   SER  142B    -35.891  34.617 -82.608  1.00 14.38 N
ATOM   2521  CA  SER  142B    -36.360  35.177 -81.342  1.00 13.26 C
ATOM   2522  CB  SER  142B    -36.788  34.057 -80.393  1.00 12.76 C
ATOM   2523  OG  SER  142B    -35.659  33.336 -79.921  1.00 13.41 O
ATOM   2524  C   SER  142B    -35.289  36.019 -80.656  1.00 13.76 C
ATOM   2525  O   SER  142B    -35.603  36.959 -79.921  1.00 12.08 O
ATOM   2526  N   SER  143B    -34.027  35.674 -80.895  1.00 13.86 N
ATOM   2527  CA  SER  143B    -32.908  36.389 -80.290  1.00 13.82 C
ATOM   2528  CB  SER  143B    -31.601  35.637 -80.566  1.00 15.38 C
ATOM   2529  OG  SER  143B    -30.526  36.196 -79.834  1.00 20.82 O
ATOM   2530  C   SER  143B    -32.772  37.832 -80.780  1.00 12.00 C
ATOM   2531  O   SER  143B    -32.217  38.680 -80.074  1.00  9.61 O
ATOM   2532  N   LEU  144B    -33.277  38.107 -81.981  1.00 10.29 N
ATOM   2533  CA  LEU  144B    -33.187  39.441 -82.575  1.00  8.43 C
ATOM   2534  CB  LEU  144B    -33.102  39.320 -84.102  1.00  8.85 C
ATOM   2535  CG  LEU  144B    -31.904  38.586 -84.722  1.00 10.18 C
ATOM   2536  CD1 LEU  144B    -32.102  38.480 -86.229  1.00 11.73 C
ATOM   2537  CD2 LEU  144B    -30.616  39.336 -84.412  1.00 11.75 C
ATOM   2538  C   LEU  144B    -34.339  40.389 -82.220  1.00  6.83 C
ATOM   2539  O   LEU  144B    -34.245  41.597 -82.441  1.00  7.49 O
ATOM   2540  N   VAL  145B    -35.411  39.838 -81.665  1.00  5.18 N
ATOM   2541  CA  VAL  145B    -36.599  40.604 -81.311  1.00  4.97 C
ATOM   2542  CB  VAL  145B    -37.629  39.694 -80.593  1.00  3.62 C
ATOM   2543  CG1 VAL  145B    -38.841  40.515 -80.145  1.00  3.67 C
ATOM   2544  CG2 VAL  145B    -38.066  38.583 -81.532  1.00  6.84 C
ATOM   2545  C   VAL  145B    -36.369  41.865 -80.474  1.00  4.98 C
ATOM   2546  O   VAL  145B    -36.891  42.932 -80.805  1.00  3.64 O
ATOM   2547  N   SER  146B    -35.597  41.760 -79.397  1.00  5.06 N
ATOM   2548  CA  SER  146B    -35.347  42.932 -78.560  1.00  6.97 C
ATOM   2549  CB  SER  146B    -34.431  42.574 -77.388  1.00  9.91 C
ATOM   2550  OG  SER  146B    -35.131  41.798 -76.432  1.00 12.54 O
ATOM   2551  C   SER  146B    -34.758  44.111 -79.327  1.00  8.11 C
ATOM   2552  O   SER  146B    -35.273  45.228 -79.254  1.00  8.37 O
ATOM   2553  N   GLY  147B    -33.678  43.873 -80.062  1.00  5.19 N
ATOM   2554  CA  GLY  147B    -33.073  44.954 -80.817  1.00  6.87 C
ATOM   2555  C   GLY  147B    -33.967  45.453 -81.937  1.00  4.82 C
ATOM   2556  O   GLY  147B    -33.955  46.629 -82.281  1.00  5.54 O
ATOM   2557  N   TYR  148B    -34.753  44.549 -82.504  1.00  3.86 N
ATOM   2558  CA  TYR  148B    -35.649  44.887 -83.602  1.00  2.80 C
ATOM   2559  CB  TYR  148B    -36.211  43.598 -84.208  1.00  2.05 C
```

FIGURE 9 - 42

```
ATOM   2560  CG  TYR  148B   -37.141  43.798 -85.382  1.00  3.38 C
ATOM   2561  CD1 TYR  148B   -36.642  43.973 -86.674  1.00  3.49 C
ATOM   2562  CE1 TYR  148B   -37.501  44.121 -87.760  1.00  2.93 C
ATOM   2563  CD2 TYR  148B   -38.525  43.781 -85.204  1.00  1.87 C
ATOM   2564  CE2 TYR  148B   -39.394  43.932 -86.282  1.00  1.52 C
ATOM   2565  CZ  TYR  148B   -38.877  44.099 -87.555  1.00  3.36 C
ATOM   2566  OH  TYR  148B   -39.741  44.229 -88.621  1.00  3.17 O
ATOM   2567  C   TYR  148B   -36.805  45.813 -83.215  1.00  2.27 C
ATOM   2568  O   TYR  148B   -37.066  46.793 -83.913  1.00  1.00 O
ATOM   2569  N   VAL  149B   -37.504  45.512 -82.120  1.00  1.46 N
ATOM   2570  CA  VAL  149B   -38.626  46.358 -81.724  1.00  1.59 C
ATOM   2571  CB  VAL  149B   -39.600  45.645 -80.752  1.00  3.52 C
ATOM   2572  CG1 VAL  149B   -40.096  44.345 -81.380  1.00  1.00 C
ATOM   2573  CG2 VAL  149B   -38.923  45.388 -79.412  1.00  3.43 C
ATOM   2574  C   VAL  149B   -38.166  47.660 -81.087  1.00  2.32 C
ATOM   2575  O   VAL  149B   -38.938  48.602 -80.979  1.00  1.77 O
ATOM   2576  N   LEU  150B   -36.905  47.718 -80.676  1.00  1.60 N
ATOM   2577  CA  LEU  150B   -36.389  48.942 -80.068  1.00  3.12 C
ATOM   2578  CB  LEU  150B   -35.413  48.606 -78.933  1.00  3.24 C
ATOM   2579  CG  LEU  150B   -36.091  47.862 -77.770  1.00  3.19 C
ATOM   2580  CD1 LEU  150B   -35.112  47.691 -76.626  1.00  5.08 C
ATOM   2581  CD2 LEU  150B   -37.321  48.623 -77.306  1.00  5.58 C
ATOM   2582  C   LEU  150B   -35.709  49.796 -81.127  1.00  2.77 C
ATOM   2583  O   LEU  150B   -35.108  50.824 -80.825  1.00  3.41 O
ATOM   2584  N   GLY  151B   -35.824  49.367 -82.379  1.00  2.18 N
ATOM   2585  CA  GLY  151B   -35.224  50.111 -83.467  1.00  3.00 C
ATOM   2586  C   GLY  151B   -36.236  50.922 -84.255  1.00  2.51 C
ATOM   2587  O   GLY  151B   -37.448  50.706 -84.161  1.00  1.00 O
ATOM   2588  N   ALA  152B   -35.731  51.878 -85.026  1.00  2.34 N
ATOM   2589  CA  ALA  152B   -36.576  52.729 -85.855  1.00  4.72 C
ATOM   2590  CB  ALA  152B   -35.951  54.120 -85.982  1.00  4.78 C
ATOM   2591  C   ALA  152B   -36.688  52.076 -87.225  1.00  4.69 C
ATOM   2592  O   ALA  152B   -35.683  51.877 -87.904  1.00  7.78 O
ATOM   2593  N   ALA  153B   -37.907  51.742 -87.631  1.00  5.53 N
ATOM   2594  CA  ALA  153B   -38.118  51.097 -88.921  1.00  5.80 C
ATOM   2595  CB  ALA  153B   -39.550  50.590 -89.018  1.00  4.77 C
ATOM   2596  C   ALA  153B   -37.816  52.021 -90.097  1.00  6.57 C
ATOM   2597  O   ALA  153B   -37.958  53.239 -90.000  1.00  5.67 O
ATOM   2598  N   PRO  154B   -37.366  51.450 -91.223  1.00  9.25 N
ATOM   2599  CD  PRO  154B   -37.023  50.042 -91.483  1.00 10.03 C
ATOM   2600  CA  PRO  154B   -37.070  52.284 -92.390  1.00  9.15 C
ATOM   2601  CB  PRO  154B   -36.507  51.281 -93.401  1.00 10.04 C
ATOM   2602  CG  PRO  154B   -37.107  49.979 -92.983  1.00 10.60 C
ATOM   2603  C   PRO  154B   -38.350  52.964 -92.878  1.00  9.66 C
ATOM   2604  O   PRO  154B   -39.458  52.539 -92.540  1.00  8.06 O
ATOM   2605  N   ARG  155B   -38.186  54.015 -93.670  1.00 10.12 N
ATOM   2606  CA  ARG  155B   -39.307  54.776 -94.200  1.00 14.62 C
ATOM   2607  CB  ARG  155B   -38.799  55.732 -95.279  1.00 18.36 C
ATOM   2608  CG  ARG  155B   -39.867  56.602 -95.900  1.00 24.01 C
ATOM   2609  CD  ARG  155B   -40.096  56.234 -97.357  1.00 28.91 C
ATOM   2610  NE  ARG  155B   -40.781  54.952 -97.524  1.00 32.92 N
ATOM   2611  CZ  ARG  155B   -41.153  54.460 -98.702  1.00 33.54 C
ATOM   2612  NH1 ARG  155B   -40.903  55.142 -99.807  1.00 34.84 N
ATOM   2613  NH2 ARG  155B   -41.784  53.295 -98.781  1.00 34.76 N
ATOM   2614  C   ARG  155B   -40.447  53.922 -94.757  1.00 11.53 C
ATOM   2615  O   ARG  155B   -41.610  54.127 -94.408  1.00 12.22 O
ATOM   2616  N   SER  156B   -40.114  52.964 -95.616  1.00 10.15 N
ATOM   2617  CA  SER  156B   -41.126  52.106 -96.224  1.00  8.14 C
ATOM   2618  CB  SER  156B   -40.465  51.057 -97.124  1.00  7.58 C
ATOM   2619  OG  SER  156B   -39.666  50.154 -96.379  1.00  9.39 O
ATOM   2620  C   SER  156B   -42.011  51.406 -95.200  1.00  7.78 C
```

FIGURE 9 - 43

```
ATOM   2621  O   SER  156B   -43.225  51.330 -95.365  1.00  7.67 O
ATOM   2622  N   GLU  157B   -41.399  50.889 -94.143  1.00  7.58 N
ATOM   2623  CA  GLU  157B   -42.153  50.190 -93.117  1.00  5.79 C
ATOM   2624  CB  GLU  157B   -41.197  49.401 -92.229  1.00  4.06 C
ATOM   2625  CG  GLU  157B   -40.617  48.190 -92.931  1.00  5.98 C
ATOM   2626  CD  GLU  157B   -39.546  47.496 -92.113  1.00  6.79 C
ATOM   2627  OE1 GLU  157B   -39.700  47.412 -90.874  1.00  7.67 O
ATOM   2628  OE2 GLU  157B   -38.559  47.024 -92.710  1.00  8.34 O
ATOM   2629  C   GLU  157B   -43.002  51.138 -92.283  1.00  5.80 C
ATOM   2630  O   GLU  157B   -44.121  50.808 -91.901  1.00  5.45 O
ATOM   2631  N   GLN  158B   -42.477  52.323 -92.001  1.00  5.94 N
ATOM   2632  CA  GLN  158B   -43.245  53.275 -91.217  1.00  7.29 C
ATOM   2633  CB  GLN  158B   -42.376  54.461 -90.818  1.00  8.03 C
ATOM   2634  CG  GLN  158B   -41.366  54.075 -89.751  1.00 13.11 C
ATOM   2635  CD  GLN  158B   -40.689  55.268 -89.115  1.00 14.69 C
ATOM   2636  OE1 GLN  158B   -41.304  56.324 -88.930  1.00 18.82 O
ATOM   2637  NE2 GLN  158B   -39.420  55.101 -88.755  1.00 13.99 N
ATOM   2638  C   GLN  158B   -44.477  53.733 -91.980  1.00  5.65 C
ATOM   2639  O   GLN  158B   -45.522  53.964 -91.381  1.00  4.85 O
ATOM   2640  N   GLU  159B   -44.360  53.848 -93.301  1.00  7.12 N
ATOM   2641  CA  GLU  159B   -45.499  54.254 -94.123  1.00  8.74 C
ATOM   2642  CB  GLU  159B   -45.111  54.382 -95.596  1.00 11.33 C
ATOM   2643  CG  GLU  159B   -43.942  55.299 -95.892  1.00 15.61 C
ATOM   2644  CD  GLU  159B   -43.862  55.661 -97.371  1.00 19.50 C
ATOM   2645  OE1 GLU  159B   -42.956  56.438 -97.745  1.00 21.70 O
ATOM   2646  OE2 GLU  159B   -44.712  55.173 -98.156  1.00 18.76 O
ATOM   2647  C   GLU  159B   -46.595  53.202 -94.014  1.00  8.03 C
ATOM   2648  O   GLU  159B   -47.769  53.530 -93.835  1.00  8.03 O
ATOM   2649  N   LEU  160B   -46.209  51.934 -94.141  1.00  7.01 N
ATOM   2650  CA  LEU  160B   -47.162  50.833 -94.055  1.00  6.37 C
ATOM   2651  CB  LEU  160B   -46.454  49.500 -94.338  1.00  6.74 C
ATOM   2652  CG  LEU  160B   -45.915  49.403 -95.770  1.00  8.35 C
ATOM   2653  CD1 LEU  160B   -45.150  48.105 -95.962  1.00  5.59 C
ATOM   2654  CD2 LEU  160B   -47.079  49.492 -96.751  1.00  5.33 C
ATOM   2655  C   LEU  160B   -47.844  50.794 -92.688  1.00  5.31 C
ATOM   2656  O   LEU  160B   -49.048  50.541 -92.589  1.00  3.17 O
ATOM   2657  N   LEU  161B   -47.083  51.058 -91.633  1.00  2.65 N
ATOM   2658  CA  LEU  161B   -47.661  51.052 -90.297  1.00  3.94 C
ATOM   2659  CB  LEU  161B   -46.594  51.319 -89.233  1.00  4.42 C
ATOM   2660  CG  LEU  161B   -46.784  50.716 -87.836  1.00  7.69 C
ATOM   2661  CD1 LEU  161B   -45.839  51.430 -86.884  1.00  8.56 C
ATOM   2662  CD2 LEU  161B   -48.218  50.841 -87.345  1.00  9.84 C
ATOM   2663  C   LEU  161B   -48.715  52.154 -90.243  1.00  2.84 C
ATOM   2664  O   LEU  161B   -49.825  51.929 -89.771  1.00  5.39 O
ATOM   2665  N   ASP  162B   -48.361  53.343 -90.730  1.00  5.24 N
ATOM   2666  CA  ASP  162B   -49.297  54.470 -90.736  1.00  5.75 C
ATOM   2667  CB  ASP  162B   -48.711  55.696 -91.456  1.00  6.09 C
ATOM   2668  CG  ASP  162B   -47.504  56.293 -90.749  1.00  9.13 C
ATOM   2669  OD1 ASP  162B   -47.408  56.227 -89.505  1.00  8.65 O
ATOM   2670  OD2 ASP  162B   -46.648  56.860 -91.458  1.00  8.00 O
ATOM   2671  C   ASP  162B   -50.603  54.096 -91.440  1.00  8.19 C
ATOM   2672  O   ASP  162B   -51.696  54.408 -90.955  1.00  6.77 O
ATOM   2673  N   GLU  163B   -50.475  53.448 -92.595  1.00  7.86 N
ATOM   2674  CA  GLU  163B   -51.625  53.027 -93.385  1.00  8.90 C
ATOM   2675  CB  GLU  163B   -51.148  52.350 -94.682  1.00 13.74 C
ATOM   2676  CG  GLU  163B   -50.443  53.306 -95.646  1.00 17.83 C
ATOM   2677  CD  GLU  163B   -49.594  52.603 -96.705  1.00 24.48 C
ATOM   2678  OE1 GLU  163B   -50.120  51.754 -97.456  1.00 26.96 O
ATOM   2679  OE2 GLU  163B   -48.387  52.912 -96.793  1.00 27.42 O
ATOM   2680  C   GLU  163B   -52.495  52.079 -92.569  1.00  8.54 C
ATOM   2681  O   GLU  163B   -53.718  52.222 -92.534  1.00  6.69 O
```

FIGURE 9 - 44

```
ATOM   2682  N   SER  164B   -51.850  51.128 -91.900  1.00  6.40 N
ATOM   2683  CA  SER  164B   -52.539  50.153 -91.062  1.00  4.06 C
ATOM   2684  CB  SER  164B   -51.521  49.183 -90.456  1.00  5.96 C
ATOM   2685  OG  SER  164B   -52.153  48.020 -89.957  1.00 10.66 O
ATOM   2686  C   SER  164B   -53.297  50.880 -89.946  1.00  4.01 C
ATOM   2687  O   SER  164B   -54.412  50.496 -89.581  1.00  4.23 O
ATOM   2688  N   ILE  165B   -52.682  51.930 -89.407  1.00  2.40 N
ATOM   2689  CA  ILE  165B   -53.300  52.729 -88.350  1.00  2.62 C
ATOM   2690  CB  ILE  165B   -52.288  53.737 -87.757  1.00  3.48 C
ATOM   2691  CG2 ILE  165B   -53.020  54.887 -87.073  1.00  4.52 C
ATOM   2692  CG1 ILE  165B   -51.353  53.001 -86.787  1.00  4.29 C
ATOM   2693  CD1 ILE  165B   -50.159  53.827 -86.308  1.00  3.59 C
ATOM   2694  C   ILE  165B   -54.507  53.474 -88.915  1.00  4.67 C
ATOM   2695  O   ILE  165B   -55.540  53.615 -88.253  1.00  4.73 O
ATOM   2696  N   ASP  166B   -54.375  53.953 -90.144  1.00  4.58 N
ATOM   2697  CA  ASP  166B   -55.482  54.650 -90.767  1.00  5.60 C
ATOM   2698  CB  ASP  166B   -55.065  55.231 -92.118  1.00  8.09 C
ATOM   2699  CG  ASP  166B   -56.092  56.196 -92.661  1.00 10.65 C
ATOM   2700  OD1 ASP  166B   -56.461  57.126 -91.918  1.00 10.81 O
ATOM   2701  OD2 ASP  166B   -56.530  56.028 -93.817  1.00 16.06 O
ATOM   2702  C   ASP  166B   -56.631  53.661 -90.947  1.00  6.94 C
ATOM   2703  O   ASP  166B   -57.795  54.005 -90.730  1.00  3.90 O
ATOM   2704  N   PHE  167B   -56.292  52.434 -91.346  1.00  5.50 N
ATOM   2705  CA  PHE  167B   -57.275  51.369 -91.543  1.00  7.94 C
ATOM   2706  CB  PHE  167B   -56.583  50.053 -91.920  1.00 11.88 C
ATOM   2707  CG  PHE  167B   -56.127  49.969 -93.352  1.00 15.37 C
ATOM   2708  CD1 PHE  167B   -55.297  48.930 -93.756  1.00 16.25 C
ATOM   2709  CD2 PHE  167B   -56.546  50.894 -94.299  1.00 16.58 C
ATOM   2710  CE1 PHE  167B   -54.892  48.811 -95.083  1.00 18.05 C
ATOM   2711  CE2 PHE  167B   -56.146  50.783 -95.630  1.00 18.43 C
ATOM   2712  CZ  PHE  167B   -55.320  49.742 -96.022  1.00 17.69 C
ATOM   2713  C   PHE  167B   -58.052  51.131 -90.256  1.00  5.81 C
ATOM   2714  O   PHE  167B   -59.275  51.003 -90.267  1.00  7.01 O
ATOM   2715  N   ALA  168B   -57.324  51.037 -89.148  1.00  4.56 N
ATOM   2716  CA  ALA  168B   -57.940  50.806 -87.847  1.00  4.69 C
ATOM   2717  CB  ALA  168B   -56.862  50.596 -86.788  1.00  3.98 C
ATOM   2718  C   ALA  168B   -58.859  51.962 -87.451  1.00  3.13 C
ATOM   2719  O   ALA  168B   -59.947  51.744 -86.921  1.00  2.35 O
ATOM   2720  N   LEU  169B   -58.430  53.194 -87.708  1.00  3.70 N
ATOM   2721  CA  LEU  169B   -59.253  54.349 -87.367  1.00  4.06 C
ATOM   2722  CB  LEU  169B   -58.493  55.646 -87.631  1.00  3.86 C
ATOM   2723  CG  LEU  169B   -57.397  55.939 -86.601  1.00  6.77 C
ATOM   2724  CD1 LEU  169B   -56.620  57.168 -87.019  1.00  7.96 C
ATOM   2725  CD2 LEU  169B   -58.018  56.138 -85.228  1.00  6.93 C
ATOM   2726  C   LEU  169B   -60.538  54.319 -88.179  1.00  5.26 C
ATOM   2727  O   LEU  169B   -61.565  54.857 -87.762  1.00  5.46 O
ATOM   2728  N   GLY  170B   -60.469  53.670 -89.337  1.00  6.56 N
ATOM   2729  CA  GLY  170B   -61.629  53.553 -90.200  1.00  6.37 C
ATOM   2730  C   GLY  170B   -62.757  52.756 -89.566  1.00  8.05 C
ATOM   2731  O   GLY  170B   -63.925  52.972 -89.890  1.00  9.59 O
ATOM   2732  N   VAL  171B   -62.426  51.827 -88.672  1.00  6.07 N
ATOM   2733  CA  VAL  171B   -63.460  51.036 -88.017  1.00  3.68 C
ATOM   2734  CB  VAL  171B   -63.154  49.516 -88.086  1.00  4.84 C
ATOM   2735  CG1 VAL  171B   -63.032  49.088 -89.533  1.00  5.50 C
ATOM   2736  CG2 VAL  171B   -61.879  49.187 -87.317  1.00  6.04 C
ATOM   2737  C   VAL  171B   -63.644  51.461 -86.559  1.00  3.29 C
ATOM   2738  O   VAL  171B   -64.084  50.679 -85.725  1.00  1.94 O
ATOM   2739  N   LEU  172B   -63.298  52.710 -86.258  1.00  2.10 N
ATOM   2740  CA  LEU  172B   -63.456  53.238 -84.909  1.00  1.93 C
ATOM   2741  CB  LEU  172B   -62.961  54.687 -84.838  1.00  1.44 C
ATOM   2742  CG  LEU  172B   -63.069  55.405 -83.490  1.00  2.60 C
```

FIGURE 9 - 45

```
ATOM   2743  CD1 LEU   172B     -62.275  54.660 -82.440  1.00  1.64 C
ATOM   2744  CD2 LEU   172B     -62.532  56.831 -83.620  1.00  1.75 C
ATOM   2745  C   LEU   172B     -64.920  53.185 -84.472  1.00  2.63 C
ATOM   2746  O   LEU   172B     -65.212  52.896 -83.314  1.00  3.20 O
ATOM   2747  N   PRO   173B     -65.864  53.471 -85.390  1.00  2.56 N
ATOM   2748  CD  PRO   173B     -65.737  53.897 -86.795  1.00  1.80 C
ATOM   2749  CA  PRO   173B     -67.280  53.426 -84.991  1.00  2.37 C
ATOM   2750  CB  PRO   173B     -68.020  53.683 -86.304  1.00  3.04 C
ATOM   2751  CG  PRO   173B     -67.082  54.565 -87.050  1.00  2.88 C
ATOM   2752  C   PRO   173B     -67.660  52.077 -84.377  1.00  2.86 C
ATOM   2753  O   PRO   173B     -68.306  52.019 -83.327  1.00  1.83 O
ATOM   2754  N   GLU   174B     -67.259  50.990 -85.031  1.00  3.72 N
ATOM   2755  CA  GLU   174B     -67.577  49.660 -84.524  1.00  3.38 C
ATOM   2756  CB  GLU   174B     -67.281  48.583 -85.575  1.00  5.65 C
ATOM   2757  CG  GLU   174B     -68.244  48.577 -86.753  1.00  6.38 C
ATOM   2758  CD  GLU   174B     -67.973  49.684 -87.757  1.00  7.10 C
ATOM   2759  OE1 GLU   174B     -68.807  49.866 -88.671  1.00  5.12 O
ATOM   2760  OE2 GLU   174B     -66.933  50.368 -87.642  1.00  3.69 O
ATOM   2761  C   GLU   174B     -66.814  49.340 -83.244  1.00  4.35 C
ATOM   2762  O   GLU   174B     -67.377  48.778 -82.302  1.00  5.56 O
ATOM   2763  N   MET   175B     -65.533  49.691 -83.200  1.00  5.24 N
ATOM   2764  CA  MET   175B     -64.764  49.411 -82.000  1.00  3.56 C
ATOM   2765  CB  MET   175B     -63.300  49.807 -82.195  1.00  3.91 C
ATOM   2766  CG  MET   175B     -62.605  49.003 -83.284  1.00  3.27 C
ATOM   2767  SD  MET   175B     -60.815  49.165 -83.218  1.00  4.52 S
ATOM   2768  CE  MET   175B     -60.632  50.873 -83.707  1.00  3.12 C
ATOM   2769  C   MET   175B     -65.373  50.151 -80.809  1.00  3.88 C
ATOM   2770  O   MET   175B     -65.578  49.566 -79.748  1.00  2.06 O
ATOM   2771  N   LEU   176B     -65.688  51.432 -80.986  1.00  3.16 N
ATOM   2772  CA  LEU   176B     -66.267  52.187 -79.888  1.00  2.84 C
ATOM   2773  CB  LEU   176B     -66.524  53.647 -80.299  1.00  4.21 C
ATOM   2774  CG  LEU   176B     -65.291  54.540 -80.494  1.00  3.57 C
ATOM   2775  CD1 LEU   176B     -65.743  55.927 -80.919  1.00  3.25 C
ATOM   2776  CD2 LEU   176B     -64.488  54.629 -79.202  1.00  1.80 C
ATOM   2777  C   LEU   176B     -67.561  51.542 -79.413  1.00  2.48 C
ATOM   2778  O   LEU   176B     -67.862  51.537 -78.223  1.00  3.56 O
ATOM   2779  N   ALA   177B     -68.315  50.976 -80.347  1.00  2.70 N
ATOM   2780  CA  ALA   177B     -69.581  50.343 -80.008  1.00  4.47 C
ATOM   2781  CB  ALA   177B     -70.474  50.281 -81.234  1.00  3.49 C
ATOM   2782  C   ALA   177B     -69.393  48.949 -79.438  1.00  4.98 C
ATOM   2783  O   ALA   177B     -70.305  48.398 -78.818  1.00  7.35 O
ATOM   2784  N   GLY   178B     -68.212  48.380 -79.649  1.00  6.22 N
ATOM   2785  CA  GLY   178B     -67.955  47.039 -79.166  1.00  6.15 C
ATOM   2786  C   GLY   178B     -68.391  46.001 -80.184  1.00  6.08 C
ATOM   2787  O   GLY   178B     -68.512  44.817 -79.859  1.00  4.07 O
ATOM   2788  N   ASP   179B     -68.633  46.441 -81.416  1.00  4.63 N
ATOM   2789  CA  ASP   179B     -69.046  45.532 -82.481  1.00  6.11 C
ATOM   2790  CB  ASP   179B     -69.922  46.277 -83.500  1.00  6.31 C
ATOM   2791  CG  ASP   179B     -70.675  45.337 -84.427  1.00  7.98 C
ATOM   2792  OD1 ASP   179B     -71.540  45.820 -85.188  1.00  7.79 O
ATOM   2793  OD2 ASP   179B     -70.403  44.119 -84.402  1.00  8.57 O
ATOM   2794  C   ASP   179B     -67.796  44.960 -83.146  1.00  6.31 C
ATOM   2795  O   ASP   179B     -67.450  45.303 -84.280  1.00  3.63 O
ATOM   2796  N   TRP   180B     -67.124  44.070 -82.425  1.00  7.68 N
ATOM   2797  CA  TRP   180B     -65.896  43.459 -82.909  1.00  8.01 C
ATOM   2798  CB  TRP   180B     -65.261  42.645 -81.781  1.00  7.72 C
ATOM   2799  CG  TRP   180B     -65.109  43.469 -80.535  1.00  6.32 C
ATOM   2800  CD2 TRP   180B     -64.208  44.564 -80.338  1.00  5.07 C
ATOM   2801  CE2 TRP   180B     -64.452  45.077 -79.047  1.00  3.64 C
ATOM   2802  CE3 TRP   180B     -63.217  45.162 -81.130  1.00  5.34 C
ATOM   2803  CD1 TRP   180B     -65.842  43.366 -79.390  1.00  5.46 C
```

FIGURE 9 - 46

```
ATOM   2804  NE1 TRP   180B    -65.456  44.328 -78.492  1.00  4.66 N
ATOM   2805  CZ2 TRP   180B    -63.742  46.162 -78.524  1.00  4.11 C
ATOM   2806  CZ3 TRP   180B    -62.509  46.245 -80.612  1.00  6.60 C
ATOM   2807  CH2 TRP   180B    -62.776  46.732 -79.319  1.00  5.80 C
ATOM   2808  C   TRP   180B    -66.071  42.605 -84.162  1.00  8.85 C
ATOM   2809  O   TRP   180B    -65.218  42.624 -85.051  1.00  7.56 O
ATOM   2810  N   THR   181B    -67.166  41.857 -84.240  1.00 10.79 N
ATOM   2811  CA  THR   181B    -67.412  41.029 -85.418  1.00 12.35 C
ATOM   2812  CB  THR   181B    -68.793  40.355 -85.353  1.00 13.87 C
ATOM   2813  OG1 THR   181B    -69.810  41.359 -85.272  1.00 18.61 O
ATOM   2814  CG2 THR   181B    -68.888  39.459 -84.136  1.00 13.72 C
ATOM   2815  C   THR   181B    -67.355  41.912 -86.663  1.00 10.61 C
ATOM   2816  O   THR   181B    -66.607  41.633 -87.600  1.00  9.89 O
ATOM   2817  N   ARG   182B    -68.139  42.989 -86.654  1.00 11.80 N
ATOM   2818  CA  ARG   182B    -68.188  43.934 -87.768  1.00 11.96 C
ATOM   2819  CB  ARG   182B    -69.237  45.010 -87.490  1.00 18.03 C
ATOM   2820  CG  ARG   182B    -70.312  45.158 -88.552  1.00 25.42 C
ATOM   2821  CD  ARG   182B    -69.759  45.626 -89.887  1.00 31.52 C
ATOM   2822  NE  ARG   182B    -70.822  46.173 -90.729  1.00 38.72 N
ATOM   2823  CZ  ARG   182B    -70.653  46.617 -91.971  1.00 41.14 C
ATOM   2824  NH1 ARG   182B    -71.686  47.099 -92.648  1.00 43.26 N
ATOM   2825  NH2 ARG   182B    -69.458  46.573 -92.543  1.00 42.09 N
ATOM   2826  C   ARG   182B    -66.830  44.602 -87.972  1.00 11.16 C
ATOM   2827  O   ARG   182B    -66.341  44.723 -89.097  1.00  9.95 O
ATOM   2828  N   ALA   183B    -66.231  45.050 -86.876  1.00  9.28 N
ATOM   2829  CA  ALA   183B    -64.929  45.699 -86.943  1.00  8.00 C
ATOM   2830  CB  ALA   183B    -64.436  46.029 -85.539  1.00  8.10 C
ATOM   2831  C   ALA   183B    -63.939  44.774 -87.649  1.00  6.64 C
ATOM   2832  O   ALA   183B    -63.239  45.185 -88.572  1.00  5.70 O
ATOM   2833  N   MET   184B    -63.900  43.518 -87.217  1.00  6.48 N
ATOM   2834  CA  MET   184B    -62.998  42.540 -87.806  1.00  9.92 C
ATOM   2835  CB  MET   184B    -63.043  41.236 -87.009  1.00 12.68 C
ATOM   2836  CG  MET   184B    -62.477  41.368 -85.604  1.00 14.45 C
ATOM   2837  SD  MET   184B    -62.275  39.781 -84.796  1.00 16.52 S
ATOM   2838  CE  MET   184B    -63.622  39.811 -83.645  1.00 16.57 C
ATOM   2839  C   MET   184B    -63.291  42.258 -89.279  1.00 10.76 C
ATOM   2840  O   MET   184B    -62.372  42.120 -90.082  1.00  9.93 O
ATOM   2841  N   GLN   185B    -64.567  42.169 -89.638  1.00 11.24 N
ATOM   2842  CA  GLN   185B    -64.916  41.902 -91.027  1.00 13.68 C
ATOM   2843  CB  GLN   185B    -66.420  41.643 -91.157  1.00 17.05 C
ATOM   2844  CG  GLN   185B    -66.911  40.452 -90.349  1.00 24.37 C
ATOM   2845  CD  GLN   185B    -68.421  40.341 -90.330  1.00 28.17 C
ATOM   2846  OE1 GLN   185B    -69.120  41.299 -89.998  1.00 30.13 O
ATOM   2847  NE2 GLN   185B    -68.933  39.166 -90.678  1.00 30.64 N
ATOM   2848  C   GLN   185B    -64.499  43.071 -91.918  1.00 11.65 C
ATOM   2849  O   GLN   185B    -64.149  42.881 -93.080  1.00 11.10 O
ATOM   2850  N   LYS   186B    -64.522  44.282 -91.370  1.00 11.72 N
ATOM   2851  CA  LYS   186B    -64.143  45.456 -92.146  1.00 10.73 C
ATOM   2852  CB  LYS   186B    -64.859  46.702 -91.606  1.00 10.52 C
ATOM   2853  CG  LYS   186B    -66.388  46.574 -91.661  1.00 11.83 C
ATOM   2854  CD  LYS   186B    -67.120  47.732 -91.001  1.00 13.74 C
ATOM   2855  CE  LYS   186B    -66.983  49.016 -91.796  1.00 15.73 C
ATOM   2856  NZ  LYS   186B    -68.014  50.019 -91.394  1.00 15.02 N
ATOM   2857  C   LYS   186B    -62.638  45.676 -92.153  1.00 12.11 C
ATOM   2858  O   LYS   186B    -62.060  45.974 -93.197  1.00  9.46 O
ATOM   2859  N   LEU   187B    -62.005  45.514 -90.992  1.00 12.40 N
ATOM   2860  CA  LEU   187B    -60.562  45.715 -90.875  1.00 15.17 C
ATOM   2861  CB  LEU   187B    -60.158  45.798 -89.400  1.00 13.66 C
ATOM   2862  CG  LEU   187B    -58.670  45.952 -89.071  1.00 11.40 C
ATOM   2863  CD1 LEU   187B    -58.124  47.248 -89.657  1.00  8.92 C
ATOM   2864  CD2 LEU   187B    -58.488  45.937 -87.560  1.00  9.50 C
```

FIGURE 9 - 47

```
ATOM   2865  C   LEU  187B    -59.741  44.627 -91.558  1.00 17.19 C
ATOM   2866  O   LEU  187B    -58.842  44.923 -92.341  1.00 18.70 O
ATOM   2867  N   HIS  188B    -60.052  43.371 -91.258  1.00 20.36 N
ATOM   2868  CA  HIS  188B    -59.324  42.245 -91.834  1.00 22.73 C
ATOM   2869  CB  HIS  188B    -59.772  40.942 -91.169  1.00 20.87 C
ATOM   2870  CG  HIS  188B    -59.606  40.934 -89.681  1.00 17.24 C
ATOM   2871  CD2 HIS  188B    -59.933  40.004 -88.753  1.00 15.71 C
ATOM   2872  ND1 HIS  188B    -59.049  41.987 -88.988  1.00 15.20 N
ATOM   2873  CE1 HIS  188B    -59.040  41.706 -87.697  1.00 13.57 C
ATOM   2874  NE2 HIS  188B    -59.571  40.509 -87.528  1.00 15.21 N
ATOM   2875  C   HIS  188B    -59.483  42.124 -93.349  1.00 26.46 C
ATOM   2876  O   HIS  188B    -58.675  41.473 -94.011  1.00 25.55 O
ATOM   2877  N   SER  189B    -60.521  42.747 -93.899  1.00 30.08 N
ATOM   2878  CA  SER  189B    -60.756  42.685 -95.337  1.00 34.21 C
ATOM   2879  CB  SER  189B    -62.258  42.706 -95.637  1.00 34.81 C
ATOM   2880  OG  SER  189B    -62.820  43.976 -95.354  1.00 35.67 O
ATOM   2881  C   SER  189B    -60.087  43.841 -96.072  1.00 36.23 C
ATOM   2882  O   SER  189B    -60.085  43.881 -97.299  1.00 36.65 O
ATOM   2883  N   GLN  190B    -59.525  44.782 -95.320  1.00 38.28 N
ATOM   2884  CA  GLN  190B    -58.861  45.936 -95.916  1.00 40.77 C
ATOM   2885  CB  GLN  190B    -58.387  46.897 -94.826  1.00 40.01 C
ATOM   2886  CG  GLN  190B    -59.519  47.554 -94.066  1.00 40.20 C
ATOM   2887  CD  GLN  190B    -60.461  48.302 -94.983  1.00 40.02 C
ATOM   2888  OE1 GLN  190B    -60.051  49.217 -95.694  1.00 40.21 O
ATOM   2889  NE2 GLN  190B    -61.732  47.913 -94.976  1.00 39.81 N
ATOM   2890  C   GLN  190B    -57.682  45.536 -96.792  1.00 42.89 C
ATOM   2891  O   GLN  190B    -57.800  45.476 -98.015  1.00 43.99 O
ATOM   2892  N   LYS  191B    -56.541  45.268 -96.170  1.00 44.78 N
ATOM   2893  CA  LYS  191B    -55.359  44.881 -96.924  1.00 47.01 C
ATOM   2894  CB  LYS  191B    -54.124  44.905 -96.017  1.00 46.09 C
ATOM   2895  CG  LYS  191B    -52.805  44.987 -96.771  1.00 44.63 C
ATOM   2896  CD  LYS  191B    -52.739  46.261 -97.605  1.00 43.97 C
ATOM   2897  CE  LYS  191B    -51.475  46.315 -98.446  1.00 43.98 C
ATOM   2898  NZ  LYS  191B    -51.437  47.531 -99.304  1.00 43.71 N
ATOM   2899  C   LYS  191B    -55.572  43.481 -97.497  1.00 49.43 C
ATOM   2900  O   LYS  191B    -56.295  42.669 -96.916  1.00 49.51 O
ATOM   2901  N   ALA  192B    -54.950  43.209 -98.642  1.00 51.33 N
ATOM   2902  CA  ALA  192B    -55.065  41.908 -99.299  1.00 52.55 C
ATOM   2903  CB  ALA  192B    -54.435  40.825 -98.425  1.00 52.62 C
ATOM   2904  C   ALA  192B    -56.520  41.553 -99.607  1.00 53.27 C
ATOM   2905  O   ALA  192B    -57.228  41.007 -98.759  1.00 53.50 O
ATOM   2906  N   GLN  193B    -56.960  41.860-100.826  1.00 53.92 N
ATOM   2907  CA  GLN  193B    -58.332  41.578-101.238  1.00 53.65 C
ATOM   2908  CB  GLN  193B    -59.161  42.864-101.212  1.00 52.88 C
ATOM   2909  CG  GLN  193B    -58.997  43.674 -99.942  1.00 51.74 C
ATOM   2910  CD  GLN  193B    -59.889  44.898 -99.907  1.00 50.95 C
ATOM   2911  OE1 GLN  193B    -61.109  44.786 -99.801  1.00 50.27 O
ATOM   2912  NE2 GLN  193B    -59.282  46.076 -99.998  1.00 49.83 N
ATOM   2913  C   GLN  193B    -58.382  40.972-102.638  1.00 54.08 C
ATOM   2914  O   GLN  193B    -58.777  39.793-102.754  1.00 54.47 O
ATOM   2915  OXT GLN  193B    -58.025  41.685-103.600  1.00 54.16 O
TER    2916       GLN  193B
ATOM   2917  OH2 WAT  500     -28.874  65.919 -68.085  1.00  2.43 O
ATOM   2918  OH2 WAT  501     -67.611  53.748 -76.569  1.00  1.89 O
ATOM   2919  OH2 WAT  502     -44.516  39.409 -89.610  1.00  5.41 O
ATOM   2920  OH2 WAT  503     -48.292  87.130 -65.406  1.00  2.69 O
ATOM   2921  OH2 WAT  504     -43.342  77.146 -82.925  1.00  4.45 O
ATOM   2922  OH2 WAT  505     -43.627  42.337 -68.612  1.00  2.91 O
ATOM   2923  OH2 WAT  506     -41.774  50.201 -76.727  1.00 22.42 O
ATOM   2924  OH2 WAT  507     -39.341  47.715 -74.095  1.00  2.10 O
ATOM   2925  OH2 WAT  508     -49.298  58.247 -58.763  1.00  4.60 O
```

FIGURE 9 - 48

```
ATOM   2926  OH2 WAT   509     -32.529  64.482 -64.542  1.00  1.00 O
ATOM   2927  OH2 WAT   510     -33.551  64.045 -55.799  1.00 11.81 O
ATOM   2928  OH2 WAT   511     -47.329  47.098 -78.494  1.00 23.19 O
ATOM   2929  OH2 WAT   512     -24.429  74.395 -52.634  1.00  1.00 O
ATOM   2930  OH2 WAT   513     -29.444  83.554 -63.396  1.00 19.77 O
ATOM   2931  OH2 WAT   514     -40.652  73.262 -84.245  1.00  4.33 O
ATOM   2932  OH2 WAT   515     -48.873  65.717 -78.845  1.00 12.28 O
ATOM   2933  OH2 WAT   516     -48.891  47.682 -89.653  1.00  5.80 O
ATOM   2934  OH2 WAT   517     -34.143  71.695 -62.997  1.00  1.00 O
ATOM   2935  OH2 WAT   518     -45.245  90.255 -70.494  1.00  7.19 O
ATOM   2936  OH2 WAT   519     -53.471  52.409 -58.930  1.00  7.74 O
ATOM   2937  OH2 WAT   520     -55.905  73.433 -67.688  1.00  7.34 O
ATOM   2938  OH2 WAT   521     -27.634  68.040 -51.809  1.00 17.79 O
ATOM   2939  OH2 WAT   522     -66.200  70.718 -71.974  1.00 24.54 O
ATOM   2940  OH2 WAT   523     -40.178  52.478 -86.168  1.00  9.81 O
ATOM   2941  OH2 WAT   524     -73.078  57.374 -86.967  1.00 21.70 O
ATOM   2942  OH2 WAT   525     -37.382  52.796 -96.658  1.00  4.47 O
ATOM   2943  OH2 WAT   526     -57.361  86.830 -65.333  1.00  3.20 O
ATOM   2944  OH2 WAT   527     -26.448  72.367 -73.520  1.00  9.73 O
ATOM   2945  OH2 WAT   528     -26.998  80.928 -80.281  1.00 11.30 O
ATOM   2946  OH2 WAT   529     -53.022  40.505 -79.031  1.00  5.42 O
ATOM   2947  OH2 WAT   530     -36.785  90.413 -73.987  1.00  7.07 O
ATOM   2948  OH2 WAT   531     -23.314  72.083 -51.179  1.00  6.51 O
ATOM   2949  OH2 WAT   532     -60.473  58.993 -85.958  1.00  5.73 O
ATOM   2950  OH2 WAT   533     -41.935  58.655 -67.926  1.00  2.01 O
ATOM   2951  OH2 WAT   534     -66.035  45.633 -75.837  1.00  8.61 O
ATOM   2952  OH2 WAT   535     -48.936  69.151 -55.259  1.00  7.03 O
ATOM   2953  OH2 WAT   536     -19.908  62.629 -62.163  1.00  4.52 O
ATOM   2954  OH2 WAT   537     -41.558  47.570 -89.150  1.00  1.70 O
ATOM   2955  OH2 WAT   538     -48.711  80.727 -54.640  1.00  5.68 O
ATOM   2956  OH2 WAT   539     -45.171  58.830 -61.103  1.00  4.30 O
ATOM   2957  OH2 WAT   540     -42.233  65.934 -78.401  1.00  6.74 O
ATOM   2958  OH2 WAT   541     -49.218  33.925 -76.362  1.00 23.24 O
ATOM   2959  OH2 WAT   542     -29.959  63.826 -42.218  1.00 31.93 O
ATOM   2960  OH2 WAT   543     -33.616  74.094 -52.418  1.00  5.03 O
ATOM   2961  OH2 WAT   544     -39.732  65.915 -68.894  1.00  5.31 O
ATOM   2962  OH2 WAT   545     -42.575  62.172 -76.510  1.00 12.04 O
ATOM   2963  OH2 WAT   546     -55.856  71.663 -75.992  1.00 31.22 O
ATOM   2964  OH2 WAT   547     -46.227  57.928 -58.792  1.00 13.63 O
ATOM   2965  OH2 WAT   548     -32.073  39.106 -77.430  1.00 22.86 O
ATOM   2966  OH2 WAT   549     -52.543  66.358 -76.164  1.00  4.12 O
ATOM   2967  OH2 WAT   550     -49.297  65.045 -63.388  1.00  8.86 O
ATOM   2968  OH2 WAT   551     -25.592  69.757 -73.599  1.00 16.87 O
ATOM   2969  OH2 WAT   552     -58.804  65.116 -72.566  1.00  8.38 O
ATOM   2970  OH2 WAT   553     -26.130  74.821 -60.958  1.00 18.71 O
ATOM   2971  OH2 WAT   554     -29.161  72.072 -47.281  1.00 15.10 O
ATOM   2972  OH2 WAT   555     -27.107  71.690 -62.070  1.00  5.03 O
ATOM   2973  OH2 WAT   556     -46.091  58.971 -63.882  1.00  5.87 O
ATOM   2974  OH2 WAT   557     -33.487  55.967 -67.951  1.00 32.89 O
ATOM   2975  OH2 WAT   558     -52.213  63.538 -80.303  1.00  8.67 O
ATOM   2976  OH2 WAT   559     -49.589  50.293 -58.619  1.00  7.20 O
ATOM   2977  OH2 WAT   560     -40.917  57.601 -76.490  1.00  6.62 O
ATOM   2978  OH2 WAT   561     -51.141  69.395 -63.223  1.00  4.76 O
ATOM   2979  OH2 WAT   562     -18.297  83.463 -60.479  1.00 14.42 O
ATOM   2980  OH2 WAT   563     -54.682  61.774 -86.468  1.00  3.38 O
ATOM   2981  OH2 WAT   564     -45.292  57.088 -65.407  1.00  3.47 O
ATOM   2982  OH2 WAT   565     -54.396  69.950 -60.834  1.00 22.96 O
ATOM   2983  OH2 WAT   566     -51.778  67.897 -80.011  1.00 29.72 O
ATOM   2984  OH2 WAT   567     -31.724  58.107 -70.898  1.00 44.86 O
ATOM   2985  OH2 WAT   568     -37.649  74.884 -81.754  1.00  5.63 O
ATOM   2986  OH2 WAT   569     -42.920  69.003 -83.056  1.00  5.50 O
```

FIGURE 9 - 49

```
ATOM   2987  OH2 WAT   570     -56.313  54.666 -59.128  1.00  5.05 O
ATOM   2988  OH2 WAT   571     -70.993  49.320 -84.273  1.00 19.64 O
ATOM   2989  OH2 WAT   572     -40.627  89.783 -66.290  1.00  5.78 O
ATOM   2990  OH2 WAT   573     -25.584  70.333 -66.683  1.00  8.26 O
ATOM   2991  OH2 WAT   574     -49.454  90.282 -53.536  1.00 23.08 O
ATOM   2992  OH2 WAT   575     -39.180  56.595 -60.721  1.00 18.76 O
ATOM   2993  OH2 WAT   576     -49.046  69.005 -61.766  1.00  3.32 O
ATOM   2994  OH2 WAT   577     -50.548  88.363 -66.826  1.00 10.15 O
ATOM   2995  OH2 WAT   578     -42.108  39.729 -90.878  1.00  5.77 O
ATOM   2996  OH2 WAT   579     -49.637  81.204 -41.199  1.00 22.48 O
ATOM   2997  OH2 WAT   580     -45.803  22.824 -67.885  1.00 24.82 O
ATOM   2998  OH2 WAT   581     -49.850  64.475 -80.842  1.00 16.19 O
ATOM   2999  OH2 WAT   582     -43.697  58.079 -85.164  1.00 22.41 O
ATOM   3000  OH2 WAT   583     -47.304  46.115 -58.797  1.00 23.89 O
ATOM   3001  OH2 WAT   584     -82.717  38.969 -74.075  1.00 34.74 O
ATOM   3002  OH2 WAT   585     -49.715  89.341 -59.296  1.00 12.64 O
ATOM   3003  OH2 WAT   586     -32.700  67.985 -53.140  1.00  4.17 O
ATOM   3004  OH2 WAT   587     -45.974  58.694 -82.154  1.00 30.94 O
ATOM   3005  OH2 WAT   588     -55.809  61.356 -93.345  1.00 13.49 O
ATOM   3006  OH2 WAT   589     -20.650  75.264 -59.362  1.00 16.77 O
ATOM   3007  OH2 WAT   590     -66.548  52.138 -90.078  1.00 14.14 O
ATOM   3008  OH2 WAT   591     -54.925  55.459 -95.849  1.00  7.81 O
ATOM   3009  OH2 WAT   592     -68.779  49.738 -71.373  1.00 10.28 O
ATOM   3010  OH2 WAT   593     -69.902  55.974 -97.961  1.00 27.85 O
ATOM   3011  OH2 WAT   594     -21.182  61.829 -57.953  1.00 23.95 O
ATOM   3012  OH2 WAT   595     -83.212  39.144 -79.545  1.00 29.82 O
ATOM   3013  OH2 WAT   596     -43.324  42.557 -94.894  1.00 24.02 O
ATOM   3014  OH2 WAT   597     -34.856 104.038 -57.220  1.00 28.10 O
ATOM   3015  OH2 WAT   598     -43.277  35.521 -82.872  1.00 16.35 O
ATOM   3016  OH2 WAT   599     -47.665  40.407 -79.317  1.00 24.22 O
ATOM   3017  OH2 WAT   600     -41.789  57.503 -91.480  1.00 18.83 O
ATOM   3018  OH2 WAT   601     -48.512  59.761 -91.146  1.00 20.86 O
ATOM   3019  OH2 WAT   602     -26.123  59.056 -60.070  1.00 11.73 O
ATOM   3020  OH2 WAT   603     -32.363  48.474 -81.838  1.00 13.98 O
ATOM   3021  OH2 WAT   604     -26.840  66.220 -42.498  1.00 25.94 O
ATOM   3022  OH2 WAT   605     -38.264  43.071 -76.420  1.00 21.75 O
ATOM   3023  OH2 WAT   606     -50.003  38.287 -79.803  1.00 42.89 O
ATOM   3024  OH2 WAT   607     -62.839  40.276 -67.265  1.00 25.59 O
ATOM   3025  OH2 WAT   608     -39.910  60.418 -53.586  1.00 20.71 O
ATOM   3026  OH2 WAT   609     -63.719  50.729 -92.938  1.00 17.10 O
ATOM   3027  OH2 WAT   610     -22.127  71.675 -40.933  1.00 28.46 O
ATOM   3028  OH2 WAT   611     -53.831  34.308 -88.924  1.00 22.50 O
ATOM   3029  OH2 WAT   612     -54.287  83.976 -60.888  1.00 23.72 O
ATOM   3030  OH2 WAT   613     -36.062  70.496 -69.379  1.00 21.98 O
ATOM   3031  OH2 WAT   614     -43.307  75.217 -46.355  1.00  6.28 O
ATOM   3032  OH2 WAT   615     -28.396  86.785 -52.663  1.00 24.20 O
ATOM   3033  OH2 WAT   616     -46.084  92.848 -53.878  1.00 20.15 O
ATOM   3034  OH2 WAT   617     -60.814  38.073 -93.015  1.00 20.74 O
ATOM   3035  OH2 WAT   618     -59.612  62.709 -68.596  1.00 36.20 O
ATOM   3036  OH2 WAT   619     -51.248  83.905 -75.376  1.00 26.85 O
ATOM   3037  OH2 WAT   620     -44.099  51.578 -98.251  1.00 13.41 O
ATOM   3038  OH2 WAT   621     -42.396  46.022 -94.960  1.00  8.24 O
ATOM   3039  OH2 WAT   622     -66.845  48.492 -67.400  1.00  4.07 O
ATOM   3040  OH2 WAT   623     -50.136  80.856 -49.996  1.00 26.60 O
ATOM   3041  OH2 WAT   624     -53.952  95.314 -73.624  1.00 21.03 O
ATOM   3042  OH2 WAT   625     -73.662  40.830 -79.214  1.00 24.94 O
ATOM   3043  OH2 WAT   626     -31.715  73.138 -79.112  1.00 23.50 O
ATOM   3044  OH2 WAT   627     -36.603  65.701 -54.556  1.00 24.30 O
ATOM   3045  OH2 WAT   628     -64.004  55.965 -69.412  1.00  6.63 O
ATOM   3046  OH2 WAT   629     -62.949  68.690 -92.871  1.00 27.96 O
ATOM   3047  OH2 WAT   630     -49.783  35.513 -83.848  1.00 20.33 O
```

FIGURE 9 - 50

```
ATOM   3048  OH2 WAT   631     -63.598  78.422 -72.835  1.00 25.41 O
ATOM   3049  OH2 WAT   632     -42.183  65.909 -51.544  1.00 21.38 O
ATOM   3050  OH2 WAT   633     -41.945  58.527 -96.501  1.00 14.82 O
ATOM   3051  OH2 WAT   634     -34.964  66.102 -42.146  1.00 15.56 O
ATOM   3052  OH2 WAT   635     -38.461  92.490 -65.765  1.00 23.14 O
ATOM   3053  OH2 WAT   636     -67.825  40.637 -77.526  1.00 23.30 O
ATOM   3054  OH2 WAT   637     -32.354  42.791 -83.903  1.00 14.98 O
ATOM   3055  OH2 WAT   638     -44.596  56.901 -73.584  1.00 14.80 O
ATOM   3056  OH2 WAT   639     -48.876  71.976 -55.608  1.00  7.03 O
ATOM   3057  OH2 WAT   640     -39.448  34.131 -88.064  1.00 11.99 O
ATOM   3058  OH2 WAT   641     -39.917  93.162 -73.439  1.00 30.97 O
ATOM   3059  OH2 WAT   642     -33.660  53.432 -92.069  1.00 31.39 O
ATOM   3060  OH2 WAT   643     -41.670  52.447 -64.837  1.00 13.10 O
ATOM   3061  OH2 WAT   644     -29.324  62.498 -45.505  1.00 12.61 O
ATOM   3062  OH2 WAT   645     -65.357  72.645 -60.451  1.00 25.51 O
ATOM   3063  OH2 WAT   646     -28.651  88.071 -79.328  1.00 24.53 O
ATOM   3064  OH2 WAT   647     -47.145  51.248 -72.486  1.00  9.60 O
ATOM   3065  OH2 WAT   648     -35.866  56.253 -67.662  1.00 21.12 O
ATOM   3066  OH2 WAT   649     -49.423  84.893 -80.250  1.00 27.48 O
ATOM   3067  OH2 WAT   650     -51.578  72.074 -76.054  1.00 20.26 O
ATOM   3068  OH2 WAT   651     -36.003  66.896 -57.822  1.00 11.00 O
ATOM   3069  OH2 WAT   652     -27.746  69.561 -68.222  1.00  6.61 O
ATOM   3070  OH2 WAT   653     -35.305  62.489 -53.683  1.00 11.33 O
ATOM   3071  OH2 WAT   654     -37.366  66.471 -69.989  1.00 35.01 O
ATOM   3072  OH2 WAT   655     -70.670  36.212 -82.244  1.00 23.57 O
ATOM   3073  OH2 WAT   656     -44.047  57.339 -78.808  1.00 22.02 O
ATOM   3074  OH2 WAT   657     -50.561  75.584 -76.121  1.00 20.22 O
ATOM   3075  OH2 WAT   658     -15.309  58.222 -44.891  1.00 25.77 O
ATOM   3076  OH2 WAT   659     -46.225  62.145 -77.318  1.00 24.71 O
ATOM   3077  OH2 WAT   660     -56.712  63.601 -62.294  1.00 14.58 O
ATOM   3078  OH2 WAT   661     -40.389  73.960 -37.229  1.00 22.30 O
ATOM   3079  OH2 WAT   662     -22.839  72.577 -63.807  1.00 18.09 O
ATOM   3080  OH2 WAT   663     -57.975  65.550 -53.913  1.00 25.92 O
ATOM   3081  OH2 WAT   664     -71.192  41.260 -59.954  1.00 28.67 O
ATOM   3082  OH2 WAT   665     -42.656  89.894 -77.859  1.00 24.66 O
ATOM   3083  OH2 WAT   666     -45.603  54.220 -88.521  1.00 13.39 O
ATOM   3084  OH2 WAT   667     -45.337  69.613 -50.352  1.00 16.39 O
ATOM   3085  OH2 WAT   668     -34.536  37.234 -91.439  1.00 27.04 O
ATOM   3086  OH2 WAT   669     -45.247  60.927 -67.533  1.00  4.97 O
ATOM   3087  OH2 WAT   670     -51.808  57.150 -89.675  1.00 14.11 O
ATOM   3088  OH2 WAT   671     -16.807  66.275 -65.350  1.00 24.46 O
ATOM   3089  OH2 WAT   672     -16.261  72.537 -57.458  1.00 15.98 O
ATOM   3090  OH2 WAT   673     -42.178  70.005 -53.929  1.00  7.40 O
ATOM   3091  OH2 WAT   674     -52.244  94.415 -53.801  1.00 22.94 O
ATOM   3092  OH2 WAT   675     -32.619  51.321 -76.288  1.00 37.06 O
ATOM   3093  OH2 WAT   676     -45.403  48.572-101.205  1.00 25.47 O
ATOM   3094  OH2 WAT   677     -52.111  66.078 -62.228  1.00 24.34 O
ATOM   3095  OH2 WAT   678     -61.750  69.333 -79.347  1.00 32.32 O
ATOM   3096  OH2 WAT   679     -27.179  66.842 -85.286  1.00 13.73 O
ATOM   3097  OH2 WAT   680     -40.470  78.170 -36.171  1.00 25.27 O
ATOM   3098  OH2 WAT   681     -47.219  34.290 -68.863  1.00 23.65 O
ATOM   3099  OH2 WAT   682     -72.715  48.986 -51.050  1.00 32.18 O
ATOM   3100  OH2 WAT   683     -26.828  77.224 -62.222  1.00 21.76 O
ATOM   3101  OH2 WAT   684     -41.018  86.115 -80.528  1.00 20.41 O
ATOM   3102  OH2 WAT   685     -38.365  62.179 -65.591  1.00  6.86 O
ATOM   3103  OH2 WAT   686     -41.541  62.923 -64.629  1.00 21.34 O
ATOM   3104  OH2 WAT   687     -41.815  58.895 -65.055  1.00 11.86 O
ATOM   3105  OH2 WAT   688     -40.894  62.440 -67.124  1.00 15.33 O
ATOM   3106  OH2 WAT   689     -41.485  60.765 -69.666  1.00 17.52 O
ATOM   3107  OH2 WAT   690     -41.468  61.404 -72.328  1.00 12.37 O
ATOM   3108  OH2 WAT   691     -40.102  66.180 -76.771  1.00  1.00 O
```

FIGURE 9- 51

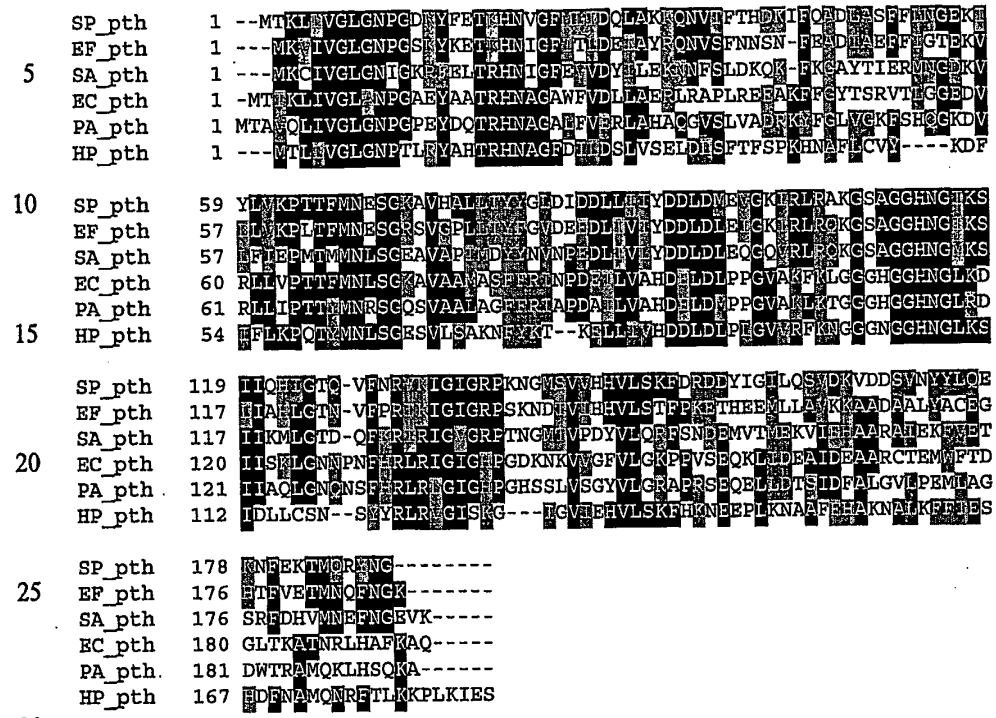
ATOM	3109	OH2	WAT	692	-39.806	62.467	-74.126	1.00	4.60 O
ATOM	3110	OH2	WAT	693	-42.980	71.805	-83.551	1.00	2.52 O
ATOM	3111	OH2	WAT	694	-41.580	65.253	-81.134	1.00	8.45 O
ATOM	3112	OH2	WAT	695	-43.917	53.865	-83.738	1.00	21.44 O
ATOM	3113	OH2	WAT	696	-44.086	56.546	-82.755	1.00	15.41 O
ATOM	3114	OH2	WAT	697	-42.412	51.507	-86.715	1.00	11.01 O
ATOM	3115	OH2	WAT	698	-43.228	59.613	-77.468	1.00	27.31 O
ATOM	3116	OH2	WAT	699	-42.404	59.741	-74.880	1.00	12.29 O
ATOM	3117	OH2	WAT	700	-46.391	57.416	-80.153	1.00	16.42 O
ATOM	3118	OH2	WAT	701	-41.089	66.800	-61.597	1.00	18.81 O
ATOM	3119	OH2	WAT	702	-37.974	67.845	-53.718	1.00	23.42 O
ATOM	3120	OH2	WAT	703	-48.700	84.329	-51.893	1.00	14.05 O
ATOM	3121	OH2	WAT	704	-46.271	87.086	-52.023	1.00	13.65 O
ATOM	3122	OH2	WAT	705	-46.661	78.833	-50.542	1.00	14.53 O
ATOM	3123	OH2	WAT	706	-42.337	83.119	-49.717	1.00	11.09 O
ATOM	3124	OH2	WAT	707	-50.418	82.903	-55.526	1.00	21.86 O
ATOM	3125	OH2	WAT	708	-29.531	84.820	-57.540	1.00	19.33 O
ATOM	3126	OH2	WAT	709	-30.116	87.453	-58.115	1.00	18.36 O
ATOM	3127	OH2	WAT	710	-30.542	89.329	-60.666	1.00	19.84 O
ATOM	3128	OH2	WAT	711	-34.996	90.611	-68.847	1.00	8.05 O
ATOM	3129	OH2	WAT	712	-54.120	66.823	-79.803	1.00	22.27 O
ATOM	3130	OH2	WAT	713	-58.324	69.413	-77.071	1.00	19.75 O
ATOM	3131	OH2	WAT	714	-48.943	62.144	-77.461	1.00	20.84 O
ATOM	3132	OH2	WAT	715	-45.323	52.226	-71.160	1.00	16.49 O
ATOM	3133	OH2	WAT	716	-51.680	50.518	-60.235	1.00	7.07 O
ATOM	3134	OH2	WAT	717	-46.126	52.624	-61.659	1.00	6.92 O
ATOM	3135	OH2	WAT	718	-47.043	46.373	-63.059	1.00	20.47 O
ATOM	3136	OH2	WAT	719	-59.718	67.072	-74.272	1.00	4.82 O
ATOM	3137	OH2	WAT	720	-57.003	50.228	-59.428	1.00	9.27 O
ATOM	3138	OH2	WAT	721	-48.843	41.749	-61.131	1.00	8.47 O
ATOM	3139	OH2	WAT	722	-43.235	57.074	-61.340	1.00	8.81 O
ATOM	3140	OH2	WAT	723	-44.806	55.424	-58.022	1.00	18.53 O
ATOM	3141	OH2	WAT	724	-43.084	61.057	-65.565	1.00	17.04 O
ATOM	3142	OH2	WAT	725	-49.577	60.623	-57.745	1.00	8.64 O
ATOM	3143	OH2	WAT	726	-51.915	63.046	-56.367	1.00	23.76 O
ATOM	3144	OH2	WAT	727	-51.541	67.717	-65.763	1.00	9.04 O
ATOM	3145	OH2	WAT	728	-53.752	73.837	-69.450	1.00	5.30 O
ATOM	3146	OH2	WAT	729	-55.969	76.184	-66.529	1.00	9.36 O
ATOM	3147	OH2	WAT	730	-57.091	70.929	-64.486	1.00	23.11 O
ATOM	3148	OH2	WAT	731	-52.101	73.157	-73.654	1.00	16.55 O
ATOM	3149	OH2	WAT	732	-52.586	75.551	-73.345	1.00	17.36 O
ATOM	3150	OH2	WAT	733	-69.750	54.881	-75.453	1.00	9.46 O
ATOM	3151	OH2	WAT	734	-72.039	53.568	-75.688	1.00	15.24 O
ATOM	3152	OH2	WAT	735	-68.329	60.662	-81.143	1.00	21.03 O
ATOM	3153	OH2	WAT	736	-65.631	63.386	-72.476	1.00	26.21 O
ATOM	3154	OH2	WAT	737	-39.896	79.834	-82.680	1.00	7.19 O
ATOM	3155	OH2	WAT	738	-39.427	83.914	-83.210	1.00	20.00 O
ATOM	3156	OH2	WAT	739	-55.916	58.914	-58.342	1.00	5.21 O
ATOM	3157	OH2	WAT	740	-62.322	49.487	-59.298	1.00	20.88 O
ATOM	3158	OH2	WAT	741	-63.468	49.112	-55.511	1.00	23.08 O
ATOM	3159	OH2	WAT	742	-58.861	62.061	-61.797	1.00	18.02 O
ATOM	3160	OH2	WAT	743	-58.819	65.975	-69.609	1.00	24.05 O
ATOM	3161	OH2	WAT	744	-62.183	57.343	-87.126	1.00	7.54 O
ATOM	3162	OH2	WAT	745	-70.010	53.666	-82.601	1.00	10.37 O
ATOM	3163	OH2	WAT	746	-69.877	56.234	-84.527	1.00	12.50 O
ATOM	3164	OH2	WAT	747	-72.977	49.645	-78.601	1.00	22.57 O
ATOM	3165	OH2	WAT	748	-69.175	41.397	-81.847	1.00	10.89 O
ATOM	3166	OH2	WAT	749	-61.141	50.911	-92.129	1.00	10.91 O
ATOM	3167	OH2	WAT	750	-65.371	55.486	-89.978	1.00	20.17 O
ATOM	3168	OH2	WAT	751	-44.575	94.915	-54.555	1.00	15.35 O
ATOM	3169	OH2	WAT	752	-69.165	43.817	-77.156	1.00	4.47 O

FIGURE 9 - 52

```
ATOM   3170  OH2 WAT   753     -55.982  38.887 -84.625  1.00 15.37 O
ATOM   3171  OH2 WAT   754     -60.493  37.315 -81.707  1.00 12.92 O
ATOM   3172  OH2 WAT   755     -43.223  49.900 -88.973  1.00  6.43 O
ATOM   3173  OH2 WAT   756     -55.017  45.768 -92.019  1.00 27.73 O
ATOM   3174  OH2 WAT   757     -44.952  44.733 -95.044  1.00 16.18 O
ATOM   3175  OH2 WAT   758     -43.718  43.045 -80.231  1.00  5.67 O
ATOM   3176  OH2 WAT   759     -45.942  40.566 -81.418  1.00 14.60 O
ATOM   3177  OH2 WAT   760     -45.034  41.477 -66.281  1.00 12.84 O
ATOM   3178  OH2 WAT   761     -51.550  38.694 -61.458  1.00 19.53 O
ATOM   3179  OH2 WAT   762     -42.740  62.341 -84.444  1.00 21.85 O
ATOM   3180  OH2 WAT   763     -56.049  52.138 -58.025  1.00  7.41 O
ATOM   3181  OH2 WAT   764     -51.437  87.770 -57.432  1.00  7.14 O
ATOM   3182  OH2 WAT   765     -50.894  84.776 -57.652  1.00 12.04 O
ATOM   3183  OH2 WAT   766     -41.068  96.729 -56.294  1.00 13.19 O
ATOM   3184  OH2 WAT   767     -42.400  41.975 -92.549  1.00 11.88 O
ATOM   3185  OH2 WAT   768     -38.040  41.495 -91.375  1.00 21.71 O
ATOM   3186  OH2 WAT   769     -34.544  47.462 -88.689  1.00 21.51 O
ATOM   3187  OH2 WAT   770     -38.510  46.508 -95.240  1.00 17.21 O
ATOM   3188  OH2 WAT   771     -37.184  49.644 -97.342  1.00 17.56 O
ATOM   3189  OH2 WAT   772       2.523   3.474  -1.217  1.00 18.99 O
ATOM   3190  OH2 WAT   773       0.570  -1.642  -3.049  1.00 20.90 O
ATOM   3191  OH2 WAT   774      -3.064  -2.461   5.287  1.00 19.90 O
ATOM   3192  OH2 WAT   775       1.091   0.254  -4.668  1.00 19.49 O
ATOM   3193  OH2 WAT   776       7.918  -6.547   0.362  1.00 15.05 O
ATOM   3194  OH2 WAT   777       0.233   1.951   4.212  1.00 23.94 O
ATOM   3195  OH2 WAT   778       3.961  -4.043   6.036  1.00 16.59 O
ATOM   3196  OH2 WAT   779       2.100  -3.174  -8.422  1.00 14.79 O
ATOM   3197  OH2 WAT   780      -2.941   0.835  -5.354  1.00 16.24 O
ATOM   3198  OH2 WAT   781      -0.585   7.197   2.315  1.00 18.00 O
ATOM   3199  OH2 WAT   782      -2.156   0.639  -2.928  1.00 17.79 O
ATOM   3200  OH2 WAT   783      -0.385   5.861  -0.531  1.00 22.14 O
ATOM   3201  OH2 WAT   784       6.438   6.917  -5.961  1.00 24.62 O
ATOM   3202  OH2 WAT   785       5.781   5.006   3.972  1.00 18.23 O
ATOM   3203  OH2 WAT   786       1.129   0.573  -2.150  1.00 19.51 O
ATOM   3204  OH2 WAT   787      -3.936   1.052   5.332  1.00 17.79 O
ATOM   3205  OH2 WAT   788       1.740  -3.827  -4.998  1.00 21.02 O
ATOM   3206  OH2 WAT   789      -0.586   1.220   1.915  1.00 20.74 O
ATOM   3207  OH2 WAT   790       1.555   3.322   8.616  1.00 13.94 O
ATOM   3208  OH2 WAT   791      -1.139   3.378  -1.612  1.00 20.00 O
ATOM   3209  OH2 WAT   792       5.372  -1.790  -4.262  1.00 19.68 O
ATOM   3210  OH2 WAT   793       0.478  -4.223  -2.400  1.00 17.40 O
ATOM   3211  OH2 WAT   794      -1.552  -1.533  -1.565  1.00 21.79 O
ATOM   3212  OH2 WAT   795      -2.043  -3.685  -0.805  1.00 16.20 O
ATOM   3213  OH2 WAT   796       1.327   2.105   0.154  1.00 26.24 O
ATOM   3214  OH2 WAT   797       2.684  -1.188   6.167  1.00 21.75 O
ATOM   3215  OH2 WAT   798      -2.634  -1.015   0.687  1.00 19.78 O
ATOM   3216  OH2 WAT   799      -0.700   8.291  -1.151  1.00 20.38 O
ATOM   3217  OH2 WAT   800      -0.849  -6.679  -0.139  1.00 19.36 O
ATOM   3218  OH2 WAT   801      -4.841  -4.586   8.701  1.00 14.17 O
ATOM   3219  OH2 WAT   802      -2.730  -4.596   1.205  1.00 19.89 O
ATOM   3220  OH2 WAT   803      -4.608   2.418   3.303  1.00 22.79 O
ATOM   3221  OH2 WAT   804       1.625   6.372   4.374  1.00 17.35 O
ATOM   3222  OH2 WAT   805      -2.704  -0.316   3.073  1.00 16.90 O
ATOM   3223  OH2 WAT   806      -2.396   0.709  -7.936  1.00 17.87 O
ATOM   3224  OH2 WAT   807      -2.103   9.519  -4.937  1.00 16.83 O
ATOM   3225  OH2 WAT   808      -2.471   3.266   0.657  1.00 17.50 O
ATOM   3226  OH2 WAT   809      -2.753   6.947  -2.234  1.00 22.55 O
ATOM   3227  OH2 WAT   810      -0.871   5.046  -5.874  1.00 17.46 O
ATOM   3228  OH2 WAT   811      -0.400  -1.342  -9.477  1.00 19.07 O
ATOM   3229  OH2 WAT   812       5.057  -1.388   6.893  1.00 20.15 O
TER    3230       WAT   812
```

FIGURE 9 - 53

FIGURE 10



The species name abbreviations are as follows: SP, *Streptococcus pneumoniae*; EF, *Enterococcus faecalis*; EC, *Escherichia coli*; SA, *Staphylococcus aureus*; PA, *Pseudomonas aeruginosa*; HP, *Helicobacter pylori*.

FIGURE 11

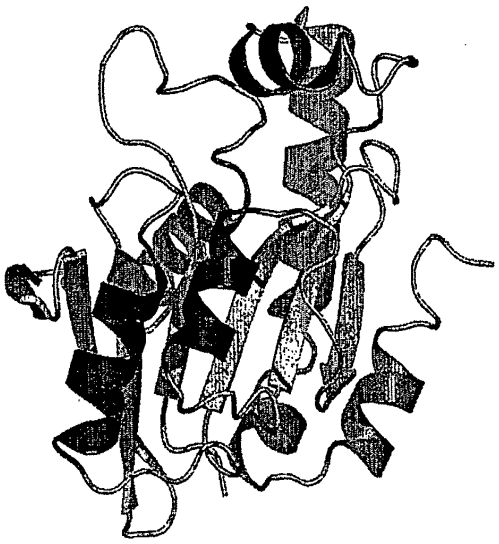


FIGURE 12

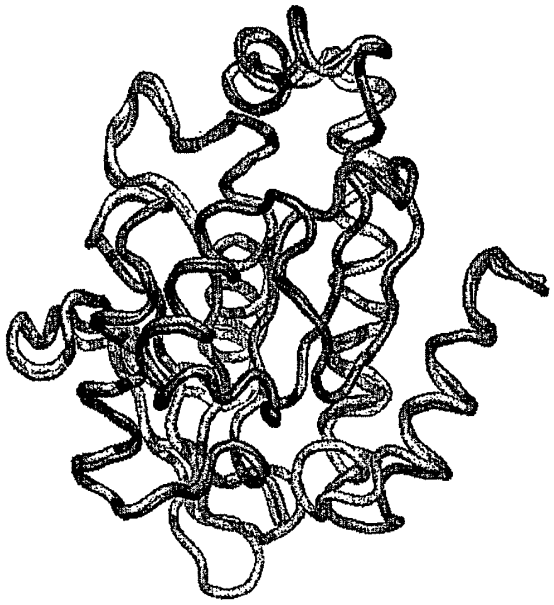
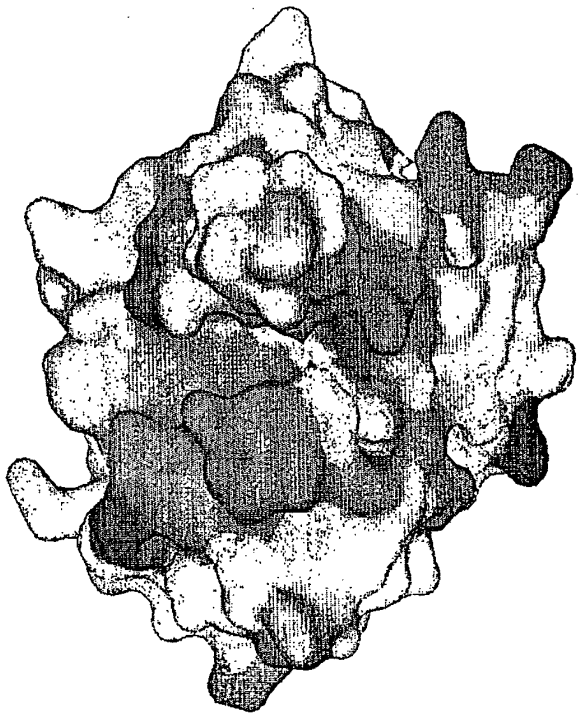


FIGURE 13

FIGURE 14



5